

75938

STIC-Biotech/ChemLib

From: Fredman, Jeffrey
Sent: Tuesday, September 17, 2002 2:36 PM
To: STIC-Biotech/ChemLib
Cc: Leffers, Gerald
Subject: FW: 09/846,456

PLEASE RUSH.

I Approve.

Jeff Fredman

CRFE

-----Original Message-----

From: Leffers, Gerald
Sent: Tuesday, September 17, 2002 2:27 PM
To: Fredman, Jeffrey
Subject: RE: 09/846,456

Hi Jeff, I'm not restricting because 1) the other sequences are all (supposedly) comprised within the whole ABC1 gene of SEQ ID NO: 1 (i.e. SEQ ID NO: 1 is a genomic gene sequence; SEQ ID NO: 3 corresponds to the first exon, SEQ ID NO: 5 to the second exon, etc.) and 2) the case has already been restricted by another examiner without restriction based upon these sequences (the case is at the stage of election by applicants). Thanks, Gerry

Gerald G. Leffers Jr., PhD
 Examiner, Art Unit 1636
 Crystal Mall 1, Room 11A09
 703-308-6232

-----Original Message-----

From: Fredman, Jeffrey
Sent: Tuesday, September 17, 2002 2:24 PM
To: Leffers, Gerald
Subject: RE: 09/846,456

Gerald,

Given this situation, why are you not restricting between the sequences and requiring election of a single DNA sequence?

Jeff

-----Original Message-----

From: Leffers, Gerald
Sent: Tuesday, September 17, 2002 2:05 PM
To: Fredman, Jeffrey
Subject: 09/846,456

Hi Jeff, please approve a RUSH search for SEQ ID NOS: 1-5 of this application. Total DNA is ~7 kb. Claims are directed to the whole sequence, or alternatively, a polynucleotide comprising 20 consecutive nucleotides of any one of SEQ ID NOS: 1-5. SEQ ID NOS: 3-5 should be comprised within SEQ ID NO: 1. I have read through the specification and have not been able to find if SEQ ID NO: 2 is found within SEQ ID NO: 1. It may be that someone could do a quick alignment of sequences in the case and determine if a single oligo search of SEQ ID NO: 1 would cover each of the other claimed sequences. Thanks, Gerry

Searcher: <u>D. Schreiber</u>	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Phone: <u>708-4292</u>	NA Sequences: <u>10</u>	STN: _____
Location: <u>CM 6A03</u>	AA Sequences: _____	DIALOG: _____
Date Picked Up: <u>9/18</u>	Structures: _____	Questel/Orbit: _____
Date Completed: <u>9/20</u>	Bibliographic: _____	DRLink: _____
Searcher Prep/Review: <u>14</u>	Litigation: _____	Lexis/Nexis: _____
Clerical: _____	Full text: _____	Sequence Sys.: <u>CompuGen</u>
Online time: <u>8</u>	Patent Family: _____	WWW/Internet: _____
	Other: _____	Other (specify): _____

24 30

12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2002, 23:21:26 ; Search time 5225.75 Seconds
(without alignments)
12938.552 Million cell updates/sec

Title: US-09-846-456-1

Perfect score: 3231

Sequence: 1 acaggcgtggtggcaggtg.....gccccacatccccaccactt 3231

Scoring table: IDENTITY_NUC

Gap 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_em.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sy.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sy.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT	1	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
AX351029	LOCUS	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
DEFINITION	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
ACCESSION	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
VERSION	AX351029.1	GI:18616385	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
KEYWORDS	AX351029.1	GI:18616385	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
SOURCE	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
ORGANISM	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
REFERENCE	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
AUTHORS	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
TITLE	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
JOURNAL	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
FEATURES	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
source	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
BASE COUNT	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002
ORIGIN	AX351029	AX351029	Sequence	1 from Patent WO0183746.	3231 bp	DNA	linear	PAT 06-FEB-2002

Query Match	100.0%;	Score 3231;	DB 6;	Length 3231;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 3231;	Conservative	0;	Mismatches	0;
1	acagggcatggtggcagggtgcctglaactcagttactcggaggtggaggtggcaatga	60		
1	ACAGGGCATGGTGGCAGGTGCCTGTAACTCAGTTACTCGGAGGTGGAGTTGCANTGA	60		
61	gccagatgcacccattgcactccagcctgggcaacaaaaggtgaaactccatccatatt	120		
61	GCCAGATGCACCAATGCACTCCAGCCTGGCGAACAAAAGTGAACATCCCATCTCAATT	120		
121	aaaaaaaaaagatgatttggctgctgcacttcaaatagtagtagaagaagaagagagag	180		
121	AAAAAAAAAAGATGATTTGGTGTGCACCTCAANTAGGTAGGAGAAGAAGAGAGAGG	180		
181	agatggagggtcaggagagatctaattactctctaaaaatcatgtagaagaataaacct	240		
181	AGATGGAGGTCAGGAGAGATCTAAATTACTCTCTAAAAATCATGCTAGGAAGATAACACCT	240		
241	tttaataacactctctgcttttataaacatcattctgccaggagctcaaaagtttcaaca	300		
241	TTTAAATAACACTCTCTGCTTTTATAACATCATCTCTGCCAGGAGCTCAAGGTTTCAACA	300		
301	aagttcacttccagaaaaaccctttggagaagacagaatatcacatctctctccatttta	360		
301	AAGTTCACTTTTCAGAAAAACCCTTTGAGAGAGACAGAAATATACATCTTCTCTCCATTTTA	360		
361	aagatgaagaaacagggccggggcacaaatggcttaagtctgtaatacccgagcacttgggagg	420		
361	AAGATGAAGAAACAGGGCCGGGCACAAATGGCTTAATGCTGTAAATCCAGCACCTTTGGGAGG	420		
421	ctgaggccagagatcgcttggactccagagtttgagacagacctgagataacatggcaaa	480		
421	CTGAGGCCAGAGATCGCTTGGCTCGAGATTTGACACCAGCCTGGATAACATGGCATAA	480		
481	acctgtctctcaaaaaataacaaaaattagatgggtgtggtggctgcacctgtgggt	540		
481	ACCTGTCTCTACAAAAAAATATACAAAAATTAGATGGGTGTGGTGCATGCACCTGTGGT	540		
541	ccagcttacttggaggctaaagttggaggatcgcttggaccagggagtcgaagttctaca	600		
541	CCCAGCTACTTGGGAGGCTAAGTGGGAGGATCGCTTGACCCAGGGAGTCAAGTCTACA	600		
601	ctgagccatgattggatcactgcactccagcctgggttagacagacaagacctgtctca	660		
601	CTGAGCCATGATTGGATCACTGCATCTCAGCCTGGGTAGACAGACAAAGACCTGTCTCA	660		
661	aaaaaagaataagaagaaagaagaagagagagagagagatgaggggaggagg	720		
661	AAAAAAGAATAAAGAGAAAGAAGAAAGAGAGAGAGAGAGATGAGGGGAGGAGG	720		
721	gaggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag	780		
721	GAGGGGGAGGAAGGAAGGAGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	780		
781	gatgaacagaggcagaagaacttaccgtaaattgctcatctgtggttgtcaagtttga	840		
781	GATGAACAGAGGCAGAAAGACTTTACGTAAATTGCTCATCATGTGTTGTCAAGTTTGA	840		
841	ccccaaacccaatttatgaccaaggttatctttgactgaggaagggggtcgctct	900		
841	CCCCAAACCCAATTATTGTGACCAAGGTTATTCTTTGACTGAGGCAAGGGGTCCGCTCT	900		
901	cttgggcttgggctttagaagctcatctctggccttcttgagatccatcccttcttt	960		
901	CCTGGGCTTGGGCTTTAGAAAGCTCATCTCTGGCCTTCTGAGATCCATCCCTTCTTT	960		
961	tatttttcttgacagcgagcttgcctgctgcactcaggctggagtcgagtgggcatgatc	1020		
961	TATTTTCTTGACAGGAGCTTGCTCTGTCACTCAGGCTGGAGTGCAGTGGCATGATC	1020		

Qy	1021	tcgactcactgtaaacctctgctcccggttccaagcgattctctctgctcgaagcctcctga	1081
Db	1021	TCGACTCAGCTGTAACTCTGCGCTCCGGGTTCAAGCGAATTCCTCTGCTCAGCTCCTCGA	1081
Qy	1081	gataacaggccgcccacacacatctgactaaattttgtatttttagtaaaagactgggtt	1140
Db	1081	GATAACAGGCGCCGCCACCACTCGGCTAATTTTGTATTTTGTAGTAAAGACTGGGTT	1140
Qy	1141	tcatactgttggccaggttggtttogaactctgacctgaggtgagctgccacccttggc	1200
Db	1141	TCATCATGTTGGCCAGGTTGGTTTGAACCTCTGACCTGAGTGTAGCTGCCACCTTGGC	1200
Qy	1201	ctcccaagtctggtgattacaggcatgacacatgcccagctcagatccatcccttt	1260
Db	1201	CTGCCAAAGTGCCTGGATTACAGCATGAGCCACATGCGCCAGCTCAGATCCATCCCTTT	1260
Qy	1261	ctaagggcaaacagtcctcattggtgcaaaaggggccatgcccaccagagttatgagtacctg	1320
Db	1261	CTAAGGCCAAACAGTCCATGGTGCAAAAGGGCCATGTCACACAGAGTTATGAGTACCTGG	1320
Qy	1321	gactccagaattccttgctgctggtggcctccacatgcaccttccagggcctgttgggcctc	1380
Db	1321	GACTCCAGAAATTCCTTGTCTGTGTGGCTCCACATGCATCTCCAGGGCCTGCTTGGGCGTC	1380
Qy	1381	ttctatgcctctcctgagctgagttgatgaaccactgatgtgagtcacctggccttgagcc	1440
Db	1381	TTCTATGCGTCTGCTCGAGTGTTGATAGAACCACTGATGTGAGTACCTGGGCTTGAGCC	1440
Qy	1441	gtggcctggagatcctctgttgactgtagcatgagggggctgtgcagctgaatgctctgca	1500
Db	1441	GTGGCCTTGGAGATCCCTTGTGACTGTAGCATGAGAGGGGCGCTGTGCAGCTGAATCTGTGA	1500
Qy	1501	tgcaggtgtggagttcttggaaatgatgaggagctggagggtgggaagagaagtaggcttg	1560
Db	1501	TGCAGGTGTGGAGTCTTGGAATATGATGGAGCTTGGAGTGGGAGAGAAAGTAGGCTTG	1560
Qy	1561	gggcagctctctcatgccacctcattctgtgcccacaaactcaggtcaaacctgtgaagagctc	1620
Db	1561	GGGCAGCTCTCTCATGCCACCTCATTTCTTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCT	1620
Qy	1621	aaatgtgaactcccttcaagtgagctacaagatattcttgcaggtaggagacctt	1680
Db	1621	AAATGTGAATCTGCCCTTCAAGTGCGCTACAAGGATCTTTGTTCAGGTAGGAGAGACTT	1680
Qy	1681	gtggcctccaagtgcacttccagggcctgcttgggcctctctctacgggtctgctcctgagt	1740
Db	1681	GTGGCCTCCACGTCGACTTCCAGGGCCTGCTTGGGCGCTCTTCTACGGGCTGTGCTCCTAGT	1740
Qy	1741	cttctatgaactcctcagggcagattcatattagaactctcacaagtttgcacctgaattt	1800
Db	1741	CTTCTATGAATCTCTCAGGCGAGATTCATATTTTACACTCTTCACAGTTTGACCTGAGTGT	1800
Qy	1801	tggccagaaataaggtgacatttagttgttgcttgaaggatgacttaaaattattagaca	1860
Db	1801	TGGCCAGAAATAAGGTGCATTTAGTTGTGGCTTGATGGATGACTTAAATATTTAGACA	1860
Qy	1861	tgtgtgtaggcctgcatctactactcttgccttttttttggccctccagtgctttgggt	1920
Db	1861	TGGTGTGTAGGCCGTGATTTCTACTCTTGCCCTTTTTTTTTTTTGGCCCTCCAGTGTGTTGGGT	1920
Qy	1921	agttttgctccctacagcgaaggcaacagagaagttggagctctggagtgctacat	1980
Db	1921	AGTTTTGCTCCCTACAGCCAAAGGCAACAGAGAAGTTGGAGTGTGGAGTGGCTACAT	1980
Qy	1981	aattttacagctgcaattctctggctgcacttcacaaatgtatacacaactaaatacaa	2040
Db	1981	AAATTTACAGACTGCAATTCCTGGCTGCATTCACAAATGTATACAAATAAATAACAA	2040
Qy	2041	gtcctgtgtttttatcacaggggaggtgatcaataataataaaaggggctggtc	2100
Db	2041	GTCTGTGTTTTTATCACAGGAGGCTGATCAATAATAATGAATTAATAAAGGGGCTGGTC	2100

Db 28640 CTCACCTGCTTTCTGCTGAGTGAAGTCACTAATAACAGAGCCCGGACCGGCGGG 28699
 Qy 2804 gaggagagagacacagggctttgaccgatagtagtaacctctgcgtcggtgcagccgaatct 2863
 Db 28700 GAGGAGGAGAGACACAGGCTTTGACCGGATAGTAACCTCTCGCTCGGTGACGCGGAATCT 28759
 Qy 2864 ataaaggaactagtcacggcaaaaccccgtaattgcagcagagagtagtgggcccg 2923
 Db 28760 ATAAAGAACTAGTCCCGCAAAACCCCGTAATTGCGAGGAGAGTAGTGAGGGCCGG 28819
 Qy 2924 gaccgcagagcagcagccctctctcccggtctgcgagcagggcagggcgagct 2983
 Db 28820 GACCGCAGAGCGAGCGACCTCTCTCCCGGCTGCGGACGCGGCGGGAGCT 28879
 Qy 2984 ccgcgcacacagagcgggtctcagggcggttctctctctctctctctctctctctctg 3043
 Db 28880 CCGCGCACCAACAGAGCGGCTTCTCAGGGCGCTTCTCTCTCTCTCTCTCTCTCTCTG 28939
 Qy 3044 ttctctccctctccggaaggtctgcaagggttaggagaaagagacgcaacacaaaa 3103
 Db 28940 TTTTCTCCCTCTTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCMACACAAA 28999
 Qy 3104 gtgaaacacaggttaagaggtctccagtgacttacttggcggttatgtttgttgcag 3163
 Db 29000 GTGAAACAGGTAAGAGGCTCTCCAGTGACTTACTTGGGCGTTATTTGTTTTCGAG 29059
 Qy 3164 gcaagagccttcggaagtctcggtttcgggactttgatccggagcccaatccc 3223
 Db 29060 GCCAAGAGGCTTCGGGAAGTGTGCTGCTTCTTGGGAGCTTTGATCCGAGCCCATCC 29119
 Qy 3224 caccactt 3231
 Db 29120 CACCATT 29127

RESULT 3
 AL359182/c
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-217B7 on chromosome 9, complete sequence.
 ACCESSION AL359182
 VERSION AL359182.20 GI:18151453
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 96717)
 Direct Submission
 Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 Cloned from request: clonerequest@sanger.ac.uk
 On Jan 15, 2002 this clone sequence replaced gi:18121468.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human

chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-217B7 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACE3.6
 IMPORTANT: This sequence is not the entire insert of clone RP11-217B7. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true right end of clone RP11-217B7 is at 96717 in this sequence. The true left end of clone RP11-122F10 is at 72980 in this sequence. The true right end of clone RP11-31J20 is at 2000 in this sequence.

FEATURES
 source
 1..96717
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-217B7"
 /clone_lib="RPCI-11.1"
 3238..3278
 /note="Sequence from AF275948 sequenced by National Institutes of Health, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA."
 84249..84273
 /note="Sequence from overlapping clone RP11-122F10 (AC026643). Assembly confirmed by restriction digest."
 92050..92163
 /note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
 92411..92557
 /note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
 BASE COUNT 27673 a 21138 c 20380 g 27526 t
 ORIGIN

Query Match 96.6%; Score 3121.6; DB 9; Length 96717;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 3207; Conservative 0; Mismatches 9; Indels 30; Gaps 5;

Qy 2 caggcgatgctgcaggtccctgtaattcagttactcgggaggtgaggttgcaatgag 61
 Db 20151 CCGGGCATGTCGCGAGTCCCTGTAATCTCAGCTACTCGGAGGTGGAGGTGCAATGAG 20092
 Qy 62 ccagatcgacacattgctactccagcctgggcaacaaagggtgaaactccatctcaatta 121
 Db 20091 CCCAGATCGCACATTGCGACTCCAGCTGGGCAACAAAAGGTGAAACTCCATCTCAATTA 20032
 Qy 122 aaaaaaaagaatgatttgggtggtcgaattcaaatagtagtagaagaagagagagga 181
 Db 20031 AAAAAAAATAATGATTTGGTGGTGCAGTTCATAATAGTAGGAGAAAGGAGAGAGA 19972
 Qy 182 gatggaggtcgaggagatcttaattactctctaaatcatgctgtaggaagataaacacct 241
 Db 19971 GATGGAGGTTCAGGAGAGATCTAATTACTCTTAATAATCATGCTAGGAAAGATAACACCTT 19912
 Qy 242 ttaataacactctctcttttataacacattctgcgaaggagctcaagggttcaacaa 301
 Db 19911 TTAATAACACTCTCTCTCTTTTATAACATCATTTCTGCAAGGAGCTCAAAAGGTTCACAA 19852
 Qy 302 agttcaatttcagaaacccctttgaggagacagacatacatctctctcattttaa 361
 Db 19851 AGTTCACTTTTCAGAAACCCCTTTGAGGAGAGACAGAAATATACATCTCTCTCATTTTAA 19792
 Qy 362 agatgaagaacagcgccggcacaattggttaatgcctgtaatccccagcactttgggagcc 421
 Db 19791 AGATGAAGAACAAGCCGGCACAATGGCTAATGCTGTAAATCCAGCAGCTTTGGGAGGC 19732
 Qy 422 tgaggccagaggatcgcttgagctccagagtttgagaccagcctgggataacatggcaaaa 481


```
TTAPVPTIMDLFQNGNWTMONPSPACSSDKIKKMLPVCPPGAGGLPPPPORKONTA
DILQDTRGNTSDLYLVKTYVQIILAKSLKNIWNEFRYGGFSLGVSNQALPPSQEVN
DAIKQMKHLKLAADSADRLNLSLGRMTGLDTRNNKVNKWNKWHAISSFLNVIN
NAILRANLQGENPSHYGITAFNHPNLTKOOLSEVALMTTSDVVLVSIQVIFAMSFV
PASVFLIQLERVSQAKHLOPLSGVKPVIYHLSNFVDMCNVVPATLVIILFICFOQ
KSYVSTNLPVIALALLLIGWSITPLMPASFVKIPSTAYVVLTSVNLFLGINGSVA
TEVLEFDNKLNNINDILKSVFLIFPHFCGLRGLDMVKNOAMADALEREGENRVS
PLSMDLVGNLFAMAVEGVVFLITVLIOYRFFIRPREVNAKLSPLNDEDDVRERO
RLDGGGNDILEIKELTKIYRRKRPAVDICVIGPPECGLLGVNAGKSTFKM
LTGDDTVTRGDAFLNKNLSILNIHEVHONMGYCPQFDATITELLTGREHVEFALLRGV
PEKEVKGWENAIRKLGVLKYGEKYAGNYSGNKRKLSANALIGGPPVPLDEPTTG
MDPKARRELWNCALSVWEGRSVVLTSMSMECECALCTRMALMVNGRRCGLSGVOHLK
NRFGGDYTVIRVAGSNPDLPKVPQDDFGLAFPGSVLKEKHRNMIQYQLPSSLSIARI
FSLQSKKRHLHIEDYSVSQTTLDQVFVNFADQSDDDHLKDLSLHKNQTVYDVAVLT
SFLODEKVKESYV"
exon      72893..72986
/gene="ABCA1"
/number=3
exon      77520..77661
/gene="ABCA1"
/number=4
exon      78931..79049
/gene="ABCA1"
/number=5
exon      101048..101169
/gene="ABCA1"
/number=6
exon      104152..104328
/gene="ABCA1"
/number=7
exon      117287..117379
/gene="ABCA1"
/number=8
exon      122329..122569
/gene="ABCA1"
/number=9
exon      125285..125424
/gene="ABCA1"
/number=10
exon      125754..125870
/gene="ABCA1"
/number=11
exon      130079..130276
/gene="ABCA1"
/number=12
exon      131023..131228
/gene="ABCA1"
/number=13
exon      131749..131925
/gene="ABCA1"
/number=14
exon      133708..133930
/gene="ABCA1"
/number=15
exon      135677..135898
/gene="ABCA1"
/number=16
exon      136959..137163
/gene="ABCA1"
/number=17
exon      138268..138381
/gene="ABCA1"
/number=18
exon      140179..140350
/gene="ABCA1"
/number=19
exon      141340..141471
/gene="ABCA1"
/number=20
exon      142777..142919
/gene="ABCA1"
/number=21
exon      143123..143260
```

```
/gene="ABCA1"
/number=22
143963..144183
/gene="ABCA1"
/number=23
145441..145513
/gene="ABCA1"
/number=24
146657..146859
/gene="ABCA1"
/number=25
148527..148575
/gene="ABCA1"
/number=26
148771..148884
/gene="ABCA1"
/number=27
150280..150428
/gene="ABCA1"
/number=28
152078..152202
/gene="ABCA1"
/number=29
153438..153536
/gene="ABCA1"
/number=30
156568..156757
/gene="ABCA1"
/number=31

Query Match      91.58; Score 2955.2; DB 9; Length 201144;
Best Local Similarity 95.18; Pred. No. 0;
Matches 3077; Conservative 0; Mismatches 121; Indels 36; Gaps 3;

Qy 2 caggcgatggtgaggtgctgtaactcagttactcggaggtgaggtgcaatgag 61
Db 31067 CCGGCATCGTGGCAGGTGCTGTAATCTCAGTTACTCGGAGGTGGAGGTGCAATGAG 31126

Qy 62 cccagatcgccaccattgcactccagcctgggcaacaaagggtgaaactccatctcaatta 121
Db 31127 CCCAGATCGCACCATTTGCATCCAGCCTCGGCAACAAAGGTGAAACTCCCATCTCAATTA 31186

Qy 122 aaaaaaaagaatgatttctgctgcacttcaaataggtaggaagaagagagagga 181
Db 31187 AAAAAAAGAAGATGATTTGGTGGTGGACTTCAATAGGTAGGAGAGAGAGAGAGGA 31246

Qy 182 gatggaggtcagggagatctaattactctctaaatcatgctaggaaagataaacacctt 241
Db 31247 GATGGAGGTCAGGAGATCTAATTACTCTCTAAATCATGCTAGGAAAGATAACACCTT 31306

Qy 242 ttaatacacctctctgtcttttataaacatctctgccaagagagctcaaaaggtttcaacaa 301
Db 31307 TTAATAACACTCTCTGCTTTTATACATCATTTGCCAAGGAGCTCAAAAGGTTTCAACAA 31366

Qy 302 agttcacttcagaaacccctttgagagacagaaatatacatctctctccattttaa 361
Db 31367 AGTTCACTTCAGAAACCCCTTTGAGGAACACAGATATACATCTCTCTCCATTTTAA 31426

Qy 362 agatgaagaacacggccgggcacaaatggcttaactcctgtaatcccagacactttggggaggc 421
Db 31427 AGATGAAGAAACAGCGCGGCACAAATGGCTAATGCTCTAATCCAGCACCTTTGGGAGGC 31486

Qy 422 tgagggccagagagatcgcttgagctccagagtttgagaccagcctggataaataagcgaacaa 481
Db 31487 TGAGGCCAGAGATCGCTTGAGCTCCAGAGTTTGAGACACGCCCTGGATTAACATGGCAAAA 31546

Qy 482 cctgtctctcaaaaaatacaaaaaattagatgggtgtgtgtggcatgacacctgtgtgc 541
Db 31547 CCCTGTCTCTACAAAAAATAACAAAAATTAGATGGGTGTGTGGCATGCACTGTGTGC 31606

Qy 542 ccagctacttggagggttaaggtgggaggaatcgcttggagccagggaggtcaagtctaacac 601
Db 31607 CCAGCTACTTGGGAGGCTAAGGTGGGAGGATCGCTTGAGCCAGGAGTCAAGTCAAGTCTACAC 31666
```


QY 602 tgagccatgattgatacactgcactccagctgggttagacagagcaagaccctgtctcaa 661
 Db 31667 TGAGCCATGATTGATGACTGCACTCCAGCTGGGTAGACAGAGCAAGACCCTGTCTCAA 31726
 QY 662 aaaaaagaaatgaaagagaaagaaagagagagagagagagagagagagagag 721
 Db 31727 AAAAAAGAAATGAAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 31786
 QY 722 aggggggag 781
 Db 31787 AGGGGGGAG 31846
 QY 782 atgaacagagggcagaaagacttaccgttaaaattgctcatcatggtgttgcaggtttgac 841
 Db 31847 ATGAACAGAGGCAGAAAGACTTTACGTAAATTGCTCATCATGTGTTCTCAAGTTTGAC 31906
 QY 842 cccaaacccaattttattgacaaaggttatcttttgaactgagggaggggtccgcctc 901
 Db 31907 CCCAAACCCCAATTTATTGACCAAGGTTATTCTTTGACTGAGCAAGGGGTCGCGCTTC 31966
 QY 902 ctgggcttgggcttttagaaagctcatctctgaccttttctgagatccatccctttttt 961
 Db 31967 CTGGGCTTGGGCTTTAGAAAGCTCATCTCTGGCCCTTTCTGAGATCCATCCCTTTCTTTT 32026
 QY 962 tatttttcttgacacggagcttctgtctctgtcactcaggtgagtgacgtggcatgatct 1021
 Db 32027 TATTTTCTTGACACGAGCTCTGCTCTGCTACTCAGCTGAGTGCACTGCGCATGATCT 32086
 QY 1022 cgaactcactgtaacctctgcctcccggttcaagagattctcctgcctcagcctcctgag 1081
 Db 32087 CGACTCAGTGTAACTCTGCCTCCCGGTTCAAGCGATTCTCCTGCTCAGCCTCCTGAG 32146
 QY 1082 ataacagggccgcacacacatctgofaattttgtatttttagtaaaagactgggttt 1141
 Db 32147 ATAACAGGGCCCGCCGACACATCTGCTGCTAAATTTTGTATTTTGTAAAGACTGGGTTT 32206
 QY 1142 catcatgttggccaggttgggttcgaactcctgacctgaggtgagctgcccaccttggcc 1201
 Db 32207 CATCATGTTGGCCAGGTGGTTTCGAACTCCTGACCTGAGGTGAGCTGCCACCTTGGCC 32266
 QY 1202 tcccaagtctgggattacaggaatgagccactgcgcccagctcagatccatcccttcc 1261
 Db 32267 TCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAGCTCAGATCCATCCCTTTC 32326
 QY 1262 taaggcacaacagtcctagtgcaaaaggccatgccacccagagttatgataactggg 1321
 Db 32327 TAAGGCAAAACAGTCCATGGTGCAAGGGGCCATGCCACCCAGAGTTATGATACCTGGG 32386
 QY 1322 actccagaattccttgccttgggtccacatgcaactccagggcctgttgggctct 1381
 Db 32387 ACTCCAGAATTCCTTGCTGGTGGCTCCACATGCATTCAGAGGCCCTGCTTGGGCCCTT 32446
 QY 1382 tctatgctctctgctgagtgatgtagaaccactgagttagtacctggggtttagccg 1441
 Db 32447 TCTATGCGTCTGTCTCTGAGTGTGATAGAACCACTGATGAGTACCTGGGCTTGGAGCG 32506
 QY 1442 tggcctggagatccttgaactgagcatggaggggcttggagctgaaatgctgcat 1501
 Db 32507 TGCCCTGGAGATCCTGTTGACTGTAGCATGGAGGGGCTTGTGCACTGAATGTCTGAT 32566
 QY 1502 gcagggtgtggagttctggaataatgtagtggaggtggagggagagagagtaggtctgg 1561
 Db 32567 GCAGGTGGTGGGATGCTGGAATATGATGAGCTGGAGTGGGAAGAGAGTAGGCTTGG 32626
 QY 1562 ggcagctctcactgccaactcatttctggccaaactcaggtcaactgtgaagagtcta 1621
 Db 32627 GGCAGCTCTCTCATGCACTCTATCTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCTA 32686
 QY 1622 aatgtgaatcctccttcaagtggtctacaaaggttatctttgtcaaggttagagaccttg 1681
 Db 32687 AATGTGAATCTGCCCTTCAAGGTGGCTCAAAAGGTATCTTTGTCAAGGTAGGAGACCTTG 32746

QY 1682 tggcctccacgtgacttccagggcctgttgggctcttctacgggtctctcctgagtc 1741
 Db 32747 TGCCCTCCACGTGCACTTCCAGGGCTGCTTGGGGCTCTTCTACGGGTCTGTCTTGAGTC 32806
 QY 1742 ttctatgaat---ccttcagggcagattcatatttagactcttccacagtttgacctgag 1797
 Db 32807 TTCTATGAATCTGCCCTTACAGGGCAGATTCATATTAGACTCTTCCACAGGGGATCAATA 32866
 QY 1798 ttitggccagaataaagtgacatttagtttggcttgatggatgaactaaattattag 1857
 Db 32867 GGTNGCCCATGCACTGGACTCAGT-----NNNAG 32895
 QY 1858 acatgggtgtaggctgcattcctactctgtcttttttttttttttttttttttttttttt 1917
 Db 32896 NCCCCCTCTCCNN 32955
 QY 1918 ggtagttttgtctccctacagccaaaggcaaacagagagagtttgaggtgttgagtgagcta 1977
 Db 32956 GGTAGTTTTGTCTCCCTACAGCCAAAGCAACAGAGAAAGTTGGAGGTCTGGAGTGGCTA 33015
 QY 1978 cataattttacagactgaattctctgtctgacctcacaaaatgatatcaaaaataaata 2037
 Db 33016 CATAATTTTACACGACTGCAATTTCTGGCTGCATTTACAAATGTATACAAACTAAATA 33075
 QY 2038 caagtctctgtttttttacagggaggtgatcaataataataataataataataataata 2097
 Db 33076 CAAGTCTCTGTTTTTATCAGGGAGGCTGATCAATATATATATATATATATATATATAT 33135
 QY 2098 gtccatattgttctgtttttttgttttttttttttttttttttttttttttttttttttt 2157
 Db 33136 GTCCATATTGTTCTGTGTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 33194
 QY 2158 gctcctctctcctcaattatgaagagaagcagtaagaatgtctcctcctcctcctgag 2217
 Db 33195 GCGCCCTCTCTCAATTTATGAGGAGAAACACCTGAGGGTTCTCTTTGGCCCTTAAA 33254
 QY 2218 ggaactcgggagctcagctgggaatctccaagggcagtagtgcgctatcaaaaatacaa 2277
 Db 33255 AGAAGTGGTACCTTAACTCGGGAATTTCCAGGCGTAATTCCTCTCATATATAAAAAA 33314
 QY 2278 gtccaggtttgtgggggaaacaaacagagcagccattaccagagagactgtccgcttcc 2337
 Db 33315 AACCAAGGTTGATCACAACAAAAACAGCAGCCATTTACCAGAGGACTGTCTGCTTCC 33374
 QY 2338 cctcaacccagccttagcctttgaaaggaaacaaagaaacaaagaaatgattggcgtcc 2397
 Db 33375 CTTACCCCGAGCTAGGCTTTGAAAGGAAACAAAGACAAAGATAATGTTGGCGTCC 33434
 QY 2398 tgaggagattcagcctagagctctctcccccactcctccctccggtgagagaaact 2457
 Db 33435 TGAGGGAGATTCAGCCTAGAGCTCTCTCTCCCCCAATCCCTCCCTCCGGCTGAGAAACT 33494
 QY 2458 aacaaagaaaaaaattgcggaaagcagagatttagaggagcaaaatccactgggtgcc 2517
 Db 33495 AACAAAGGAAAAAAATTTGCGAAAGCAGCATTTAGAGGAAGCAAAATTTCCACTGGTGCC 33554
 QY 2518 cttggctccggggaacgttgagactagagagctgcggcgagccccagagccagcgttcc 2577
 Db 33555 CTTGGCTCCGGGAACCTGGAGCTAGAGAGCTGCGGGCGAGCCCGAGCCGAGCGGCTTCC 33614
 QY 2578 cgcgctcttagc 2637
 Db 33615 CGCGGCTTTAGGCTGGCGGCCCGGGGGGAGAGGCGACGACAGCCGCGGAGCCCTAAG 33674
 QY 2638 acactgtgtacctccacccccccacccccccacccccccacccccccacccccccactagat 2697
 Db 33675 ACACCTGTGTACCTCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCTAGAT 33734
 QY 2698 gtgtgtggggcggtgaaactgtccctgttaagggggcgggccccggctccacgtgtcttc 2757
 Db 33735 GTGTGTGGGGCGGTGAACGTCGCCCGCTTTAAGGGGGGGGGCGCGGCTTCCACGTGCTTC 33794
 QY 2758 tgcgtgagtgaactacaataacagagggcggaaggggggggggagagagagagagagca 2817

```

Db 33795 TGCTAGTACTGAACATACATAACAGAGCGCGGAAGGGGGGAGGAGGAGCA 33854
Qy 2818 caggcttgaccgataagtaacctctgcgtcggtgagcggaatctataaaaggaaactag 2877
Db 33855 CAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCGGAGCGGAATCTATAAAGGAACTAG 33914
Qy 2878 tccgcgcaaaaccccgtaattcgacgacagagtagtggtggcgagcccgacagagccg 2937
Db 33915 TCTCGGCAAAACCCCGTAATTGCGAGCGAGAGTGAAGTGGGCGCGGAGCCCGAGAGCCG 33974
Qy 2938 agcgacacctctctccgggctgcgcagggcgagggcgagggcgagctcccgacccaacag 2997
Db 33975 AGCGACCCCTTCTCTCCGGGCTGCGGACGAGCGAGGCGGGAGCTCCGCGCACCAACAG 34034
Qy 2998 agcgggtctcaaggcgcttgctctctctctctctctctctctctctctctctctctct 3057
Db 34035 AGCGGTTCTCAGGGCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 34094
Qy 3058 ccggaagcctgtcaagggttaggaagagacgacgacacacacacacacacacacacaggt 3117
Db 34095 CCGGAAGGCTTGTCAAGGGGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTA 34154
Qy 3118 agagctctcagtgactacttggtggcggttattgtttgttcgagggcgagggaggttc 3177
Db 34155 AGAGGCTCTCAGTGAATCTTCTGGGGCTTATTGTTTGTTCGAGGCGCAAGGAGGCTTC 34214
Qy 3178 gggagtgctcggttctcgagacttgatccgagccacacacacacacacacacacac 3231
Db 34215 GGAAGTGCTCGGTTTCGGGACTTTGACCGGAGCGGAGCGGAGCGGAGCGGAGCGG 34268

RESULT 5
AX351031 2893 bp DNA linear PAT 06-FEB-2002
LOCUS AX351031
DEFINITION Sequence 3 from Patent WO0183746.
ACCESSION AX351031
VERSION AX351031.1 GI:18616387
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P.,
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 3 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
Location/Qualifiers
1..2893
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 745 a 681 c 769 g 698 t
ORIGIN
Query Match 89.5%; Score 2893; DB 6; Length 2893;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 acaggcgatggtgaggtgactgtaactcagttactcggaggtgaggttgcaatga 60
Db 1 ACAGGGCATGTTGGCAGTGGCTGCTTAATCTCAGTTACTCGGAGGTGAGGTTGCAATGA 60
Qy 61 gccagatcgcaacattgcactccagcctggggaacaaagggtgaactccatcaatt 120
Db 61 GCCCAGATCGCACCATTTGCACTCCAGCCTGGGCAACAAAAGGTGAAACTCCATCTCAAT 120
Qy 121 aaaaaaaagaaatgatttgggtgctgactcacaataggtaggaagagagagag 180
Db 121 AAAAAAAGAAATGATTGGTGGTGGCTGACTTCAATAGTAGGAGAGAGAGAGAGG 180

```

```

Qy 181 agatgagggtcaggagatcttaattactctctaaaaatcatgtaggaagataaacact 240
Db 181 AGATGAGGGTCAGGGAGATCTAATTACTCTCTAAAAATCATGCTAGGAAGATAAACACT 240
Qy 241 tttaataacactctctgctttttaatacatctctgcaaggagctcaaaaggtttcaaca 300
Db 241 TTTAATAACACTCTCTGCTTTTATAACATCATCTTGCACCAAGGAGCTCAAGGTTTCAACA 300
Qy 301 aagttcactttcagaaaaaccccttggagaaagacagataacatctctctcatttta 360
Db 301 AAGTTCACTTTTCAGAAAAACCCCTTTGAGAAAGACAGATATATACATCTTCTCTCCATTTA 360
Qy 361 aagatgaagaaacagcgccgggcacaaatgcttaatccctgttaatccacagcactttggagg 420
Db 361 AAGATGAAGAAACAGCGCGGGCACAAATGGCTAATGCCGTGTAATCCACAGACTTTGGGAGG 420
Qy 421 ctgagcgcaagagatcgcttgagctccagagtttgagaccagcctggataacatggcaaa 480
Db 421 CTGAGCGCCAGAGGATCGCTTGAGCTCCAGAGTTTGAGACCAGCCTGGATAACATGGCAAA 480
Qy 481 accctgtctctacaaaaaaatacaaaaaattagatgggtggtggtggtggtggtggtggt 540
Db 481 ACCCTGTCTCTACAAAAAATAACAAAAATTAGATGGGTGTTGGTGGCATGCACCTGTGTT 540
Qy 541 ccagctacttgagggttaaggtggaggatcgcttgagcccgagggaggtcaagctctaca 600
Db 541 CCAGCTACTTTGGGAGGCTAAGTGGAGGATCGCTTGAGCCACAGGAGTCAAGTCTTACA 600
Qy 601 ctgagcctatgattgatcactgcactccagcctgggtgagagagagagagagagagagag 660
Db 601 CTGAGCCATGATTGGATCACTGCCTCCAGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGT 660
Qy 661 aaaaaaagaaatgaagagagaaagaaagagagagagagagagagagagagagagagag 720
Db 661 AAAAAAAGAAATGAAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 720
Qy 721 gagggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 780
Db 721 GAGGGGGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 780
Qy 781 gatgaacagagcagagagagagagagagagagagagagagagagagagagagagagag 840
Db 781 GATGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 ccccaaaaccccaattatttgacaaaggttattcttctgactgagggcaagggggtccgctct 900
Db 841 CCCCAAAACCCCAATTATTATTGACCAAGGTATTCTTTGACTGAGGCAAGGGGTCCGCTCT 900
Qy 901 cctggccttgggctttagaagagcactctctggcctttctgagatccatcccttcttct 960
Db 901 CCTGGCCTTGGGCTTTAGAAAGCTCATCTCTGGCCTTTCTGAGATCCATCCCTTTCTTT 960
Qy 961 ttattttctgacagcagagctctctctctgactcagcagctgagtgagtgagtgagtgatc 1020
Db 961 TTATTTTTCTTGACACGGAGCTCTTGCTCTGTCTACTCAGGCTGGAGTGAGTGAGTGATC 1020
Qy 1021 tcgactcactgtaacctctgctcccggttcaagcagattctcctgctcagcctccctcga 1080
Db 1021 TCGACTCACTGTAACCTCTGCTCCCGGTTTCAAGCGATTCTCTCTGCTCAGCCCTCTGA 1080
Qy 1081 gataacagcgcccgccacacacatctggttaatttttttttttttttttttttttttttt 1140
Db 1081 GATAACAGCGCCCGCCACCATCTGGCTAATTTTTTTTATTTTATTTTATTTTATTTTATTT 1140
Qy 1141 tcatcatgttgccaggttggtttcgaactcctgactgagtgagtgagtgagtgagtgagtgag 1200
Db 1141 TCATCATGTTGGCAGGTTGGTTTTCGAACCTCTGACCTGAGTGAGTGAGTGAGTGAGTGAG 1200
Qy 1201 ctcccaaatgctgggtattacagcgatgagcgactgcgcccagctcagatccatcccttt 1260
Db 1201 CTCCCAAAAGTGTGGGATTACAGGATGAGGCATGAGCCACTGCGCCAGCTCAGATCCATCCCTTT 1260
Qy 1261 ctaaggggcaaacagtcctcattggtgcaaaagggggccaatgcccacccagaggttatagtagtacc 1320

```

```

Db 1261 CTAAGGGCAACAGTCCATGTGCAAGGGCCATGCCACCCAGATTATGAGTACCTGG 1320
QY 1321 gactcagaattcctgctggtgctcctccatgcaactccagggcctggtggcctc 1380
Db 1321 GACTCAGAAATCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1381 tctatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1440
Db 1381 TTCATAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 gggcctgagatcctgctgctgctgctgctgctgctgctgctgctgctgctg 1500
Db 1441 GTGGCTGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 tgcaggtggtgaggtcctgctgctgctgctgctgctgctgctgctgctgctg 1560
Db 1501 TGCAGGTGCTGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1561 gggcagctcctcagcagcctcctcagcagcctcctcagcagcctcctcagcag 1620
Db 1561 GGGCAGCTCTCTCATGCCACCTCATCTGCGCCAAACCTCAGGTCAAACTGTGA 1620
QY 1621 aaatgtaactgctcctcagcagcctcctcagcagcctcctcagcagcctcct 1680
Db 1621 AAATGTGAATCTGCCCTTCAAGTGGCTTCAAAAGGTATCTTGTCAAGGTAG 1680
QY 1681 gtgctcctcagcagcctcctcagcagcctcctcagcagcctcctcagcagc 1740
Db 1681 GTGGCTCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 1741 cttctataactcctcagcagcctcctcagcagcctcctcagcagcctcctcag 1800
Db 1741 CTTCTATGAATCTTCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 1801 tggcagataaggtgacatttagttgtgctgctgctgctgctgctgctgctgctg 1860
Db 1801 TGGCCAGAAATAAGGTGACATTAGTTTGTGGCTGTGATGATGACTTAAATATT 1860
QY 1861 tgggtgtagcctgctcctcagcagcctcctcagcagcctcctcagcagcct 1920
Db 1861 TGGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
QY 1921 agtttgcctcctcagcagcctcctcagcagcctcctcagcagcctcctcagc 1980
Db 1921 AGTTTGTCTCCTTACACCCAAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1981 aattttacagcagcctcctcagcagcctcctcagcagcctcctcagcagcct 2040
Db 1981 AATTTTACAGCAGCTGCAATCTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
QY 2041 gtcctgtgtttttatcacagcagcctcctcagcagcctcctcagcagcctc 2100
Db 2041 GTCCTGTGTTTTTATCACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2101 catatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2160
Db 2101 CATATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
QY 2161 tcttctcctcattatgaagaagcagcagcagcagcagcagcagcagcagcagc 2220
Db 2161 TCTTCTCTCTCAATTTATGAAGAAGCAGTGAAGATGCTCTCTCGGGTCTCTG 2220
QY 2221 cctgggagctcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2280
Db 2221 CCTGGGAGCTCAGGCTGGGAATCTTCCAAAGCAGTAGCTGCCCTATCAAAAT 2280
QY 2281 caggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2340
Db 2281 CAGGTTGTGTGGGGGAAACAAAGACAGCCCATTTACCCAGAGGAGTGTCCCT 2340
QY 2341 caccagcctagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2400

```

```

Db 2341 CACCCAGCCTAGGCCTTTGAAGGAACAAACAGACAGCAAAATCATTTGGCGCTCTGA 2400
QY 2401 gggagattcagcctagcctgctcctcctcctcctcctcctcctcctcctcctc 2460
Db 2401 GGGAGATTACGCTAGAGCTCTCTCTCCCAATCTCTCTCCGCTCGGCTAGGAACTAAC 2460
QY 2461 aaaggaaaaaaattcgaagaagcagcagcagcagcagcagcagcagcagcagc 2520
Db 2461 AAGGAAAAAAATTCGGAAAGCAGATTAGAGGAAGCAAAATCCACTGTGTGCCCTT 2520
QY 2521 ggcctcgggaacgtgactagcagcagcagcagcagcagcagcagcagcagcagc 2580
Db 2521 GGCTGCCGGAACGTGGACTTAGAGAGTCTCGGGCAGAGCCCGAGCCGCTTCCCGC 2580
QY 2581 ggccttagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2640
Db 2581 GCCTCTTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2640
QY 2641 cctgctgtacccctcagcagcagcagcagcagcagcagcagcagcagcagcagc 2700
Db 2641 CCTGCTGTACCTTCCACCCACCCACCCACCCACCCACCCACCCACCCACCTCC 2700
QY 2701 tctgtggcgctgaacgtcgcctgctgctgctgctgctgctgctgctgctgctg 2760
Db 2701 TCGTGGCGGCTGACGCTGCGCGCTTTAAGGGCGGGCGCGCGCGCGCGCTCCAC 2760
QY 2761 tgaagtactgaactacataacacagcagcagcagcagcagcagcagcagcagcag 2820
Db 2761 TGAGTACTGAACCTACATAAACAGAGCGCGGAAAGGGGGGGGAGGAGGAGAGC 2820
QY 2821 gcttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2880
Db 2821 GCCTTGACCATAGTAACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
QY 2881 cggcaaaaaacccc 2893
Db 2881 CGGCAGAAACCC 2893

```

RESULT 6

```

AF275948 149034 bp DNA linear PRI 17-JUL-2000
LOCUS Homo sapiens ABCA1 (ABCA1) gene, complete cds.
DEFINITION AF275948
ACCESSION AF275948
VERSION AF275948.1 GI:9247085
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 149034)
AUTHORS Santamarina-Fojo, S., Peterson, K., Knapper, C., Qiu, Y., Freeman, L.,
Cheng, J.F., Osorio, J., Remaley, A., Yang, X.P., Haudenschild, C.,
Prades, C., Chimini, G., Blackmon, E., Francois, T., Duverger, N.,
Rubin, E.M., Rosier, M., Benefield, P., Fredrickson, D.S. and Brewer, H.B.
Jr.
TITLE Complete genomic sequence of the human ABCA1 gene: analysis of the
human and mouse ATP-binding cassette A promoter
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
MEDLINE 20345099
REFERENCE 2 (bases 1 to 149034)
AUTHORS Santamarina-Fojo, S., Peterson, K.M., Knapper, C.L., Freeman, L.A.,
Remaley, A.T., Yang, X.-P., Haudenschild, C.C., Blackmon, E.E.,
Francois, T.L. and Brewer, H.B. Jr.
Direct Submission
Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA
FEATURES
source Location/Qualifiers
1.149034
/organism="Homo sapiens"
/db_xref="taxon:9606"
join(1454..1674,25831..25989,40385..40478,45012..45153,

```


RESULT 7
AX060715
LOCUS
DEFINITION Sequence 3 from Patent WO0078972.
ACCESSION AX060715
VERSION AX060715.1 GI:12406104
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1643)
JOURNAL Law, R.M., Wade, D. and Garvin, M.
CV Regulation with binding cassette transporter protein abc1
CV THERAPEUTICS, INC. (US) Patent: WO 0078972-A 3 28-DEC-2000;
FEATURES Location/Qualifiers
source 1..1643
BASE COUNT 370 a 413 c 457 g 403 t
ORIGIN

Query Match 46.0%; Score 1485.4; DB 6; Length 1643;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 6; Indels 26; Gaps 10;

QY 1345 gccctcacatgcactccacggcgctgtggcgctctcttatgcctgtccctgagtggt 1404
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16 GGCTCCACATGCACATCCAGGGCGTCTGG-CTCTTCTATGGTCTGTCTCGAGTGT 73

QY 1405 gatagaaccactgatgtagtagtacctggcttgagccctggcctggagatcctgttgactg 1464
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 74 GATGAAACACACTGATGTAGTACCTGGCTTTGAG-CGTGGCCTGGAGATCCTGTTGACTG 132

QY 1465 tagcatggagggggctgtgcagctgaatgtctcatgcagggtggggagttcttggaaat 1524
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 TACCATGGAGGGGGCTTGT-CAGCTGAATGTCGTATGACAGGTGGGGAGTCTGGGAAT 191

QY 1525 atgatgagctggaggtgggaagagaagtaggtgtgggcagctctctcatgccacctca 1584
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 192 ATGATGGAGCTGGAGTGGGAAGAGAGTAGGTGGGGCAGCTCTCTATGCCACTCA 251

QY 1585 ttctggccaaaactcaggtcaaaactggaagatctaaatgtaactcgcctccaagt 1644
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 252 TTCTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCTAAATGTGAATCTGCCCTTCAGGT 311

QY 1645 ggctacaaagggtatcttgtcgaagtaggagacacctgtggccctccacgtgcactccagg 1704
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 312 GGCTACAAAGGTATCTTGTCAAGTAGGAGACCTTGTGGCCTCCAGTGCACTCCAGG 371

QY 1705 gccctgttggcctcttcatcaggctctgctgaagtcttatgaat---ccttcagggc 1761
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 372 GCCCTGTT--GGCCTCTTCTACGGGTGTCTCTGAGTCTTCTATGAATCTCCCTTCAGGGC 430

QY 1762 agattcataatttagactcttcacagtttgacctgagtttggccagaaataaggtgacatt 1821
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 431 AGATTCTATTTTAGACTCTTACAGTTTGACCTGAGTTTGGCCAGAAATGAAGTGACATT 490

QY 1822 tagtttgttggttgatggatgaacttaaatatttagac--atggtgtgtaggcctgcatt 1879
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 491 TAGTTTGTGGCTTGATGAATGACCTTAAATATTTAGACATATGGTGTGAGCCTGCATT 550

QY 1880 cctactcttgccttttttttggccctccagtgtttgggtagttttgt-cctctacag 1938
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 551 CCTACTCTTGGCTTTTTTTTTTGGCCCTCCAGTGTTTGGGTAGTTTTGGTCCCCCTACAG 610

QY 1939 ccaaggcacaacagagaagtttggaggtctggagtggtacataattttcacagactgcaa 1998
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 611 CCAAGGCCAAACAGATAAGTTGGAGGTCTGGAGTGGCTACATTAATTTTACAGACTTGCAA 670

DEFINITION Sequence 3 from Patent WO0078971.
ACCESSION AX060894
VERSION AX060894.1 GI:12406271
KEYWORDS human.
ORGANISM Homo sapiens
SOURCE
REFERENCE 1 (bases 1 to 1643)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE Atp binding cassette transporter protein abc1 polypeptides
JOURNAL Patent: WO 0078971-A 3 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
1..1643
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 370 a 413 c 457 g 403 t
ORIGIN
Query Match 46.0%; Score 1485.4; DB 6; Length 1643;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 6; Indels 26; Gaps 10;
QY 1345 gctccacatgcactccagggcctgtggcctctctctatgcgtctgctgctgagtggt 1404
DB 16 GGCCTCATGATGACATCCAGGCGCTGCTGG--CTCTCTATGGGTCTGCTGAGTGT 73
QY 1405 gatagaaccacgatgtgagtgacgtggccttgagcgtggcctggagatcctgttgactg 1464
DB 74 GATAGAACCACATGATGTGAGTACCTGGGCTTGAG-CGTGGCCTGGAGATCCTGTTGACTG 132
QY 1465 tagcatgaggggcctgtgagctgaatgtctgcatgcaggtgtgtggagttctggaa 1524
DB 133 TAGCATGGAGGGGCTGTGT-CAGCTGAATGTCTGTATGCAGTGTGGGAGTCTTGGAA 191
QY 1525 atgatgagctgaggtgggaagagaagtaggttggcagctctctcatccacctca 1584
DB 192 ATGATGGAGCTGGAGTGGGAGAGAGTAGGCTTGGGCGACCTCTCTCATGCCACTCA 251
QY 1585 ttctggccaaactcaggtcgaactgtgaagagtctaaatgtgaatctgacctcaagt 1644
DB 252 TTCTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCTAAATGTGAATCTGCCCTTCAAGT 311
QY 1645 ggctacaaaggatctcttgaagtaggagacctgtggcctccacgtgcaactccagg 1704
DB 312 GGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGGCCTCCACGTGCACTTCCAGG 371
QY 1705 gctctgtggcctctctacgggtctgtcctgagctctctctatgaat---ccttcagggc 1761
DB 372 GCGTGCTT--GGCTCTCTACGGGCTGTCTCTGAGTCTCTATGAATCTCCCTTCAGGGC 430
QY 1762 agattcatatttagactcttcacgtttgacctgagtttggccagaataaagtgacatt 1821
DB 431 AGATTCTATATTAGACTCTTACAGTTTTCACCTGAGTTTGGCCAGAATAAGGTGACATT 490
QY 1822 tagtttgtgttgatgagatgaactaaatattagac--atggtgtgtaggcctgcatt 1879
DB 491 TAGTTTGTGGCTTGATGAATGACTTAAATATTTAGACATATGGTGTGAGGCTGCATT 550
QY 1880 cctactcttgctcttttttccctccagtggtttgggtagtttggct--ccctacag 1938
DB 551 CTTACTCTTGGCTTTTGTGGCTTCCAGTGTTTTGGGTAGTTTGGCTCCCTCCCTACAG 610
QY 1939 ccaaggccaaacagagaagttgaggtgtggagtggtacataaattttacagactgcaa 1998
DB 611 CCAAGGCCAAACAGATAAGTTGGAGTCTGGAGTCTGAGTACATAAATTTACAGACTGCA 670
QY 1999 ttctctggtgcacttcacaaatgtatatacaaaactaaatacaagctcgtgtttttatcac 2058
DB 671 TTCTCTGGCTGCACTTCACAAATGTATACAAACTAAATACAAAGTCTCTGTGTTTTATCAC 730

QY 2059 agggagcgtgatcaataataatgaataaaggggggtggt-ccatatgtgtctgtgttt 2117
DB 731 AGGGAGCGCTGATCAATATATATGAATTAAGAGGGGCGTGGTCCCATATATTGTTCTGTGTT 790
QY 2118 ttgtttgtgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgttt 2177
DB 791 TTG-----TTTGTGTTGTTTCTTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 841
QY 2178 tgaagagaagcagtaagaatgtctctcgtggtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2237
DB 842 TGAAGAGAAGCAGTAAGATGTTCTCTCGGGTCTCTCTGAGGAGCTGGGAGCTCAGGCT 901
QY 2238 gggatctccaaagcagtagtgcctcctatacaaaatcaaaatcaaaatcaaaatcaaaatcaaaat 2297
DB 902 GGGATCTCCAAGGCAGTAGTGCCTTATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAAT 961
QY 2298 aacaaaagcagccattaccacagaggactgtccgctcctcctcctcctcctcctcctcctcctcct 2357
DB 962 AACAAAAGCAGCCATTACCCAGAGGAGTGTCCGCTTCCCTTCCCTCAGCCAGCTTAGGCT 1021
QY 2358 ttgaaggaagaaacaaacaaacaaatgattggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2417
DB 1022 TTGAAGAGAAACAAAGACAAAGAAATGATTTGGCGTCTCTGAGGGAGATTTACGCTTGA 1081
QY 2418 gctctctcctcccaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2477
DB 1082 GCTCTCTCTCCCCANTCCCTCCCTCCGCTGAGGAACTAACAAAGGAAAGGAAAGGAAAGGAA 1141
QY 2478 cggaaagcaggtatttagaggagcaaatccactgtgtcctcgtcgtcgtcgtcgtcgtcgtcgtcgt 2537
DB 1142 CGGAAAGCAGGATTTAGAGGAAGCAAAATTCACATGCTGCTGCTTGGCTGCCGGGACGCTGG 1201
QY 2538 actagaagatctcggcgcagcccgagcccgagcccgagcccgagcccgagcccgagcccgagcccg 2597
DB 1202 ACTAGAGAGTCTGGCGCGACGCCCGGAGCCGAGCCGCGCTTCCCGCGGCTTTAGGCCGCGG 1261
QY 2598 gcccgggcggggaaggagcgcagc 2657
DB 1262 GCCCGGGCGGGGAAGGGGACGACGCGCGGACCCCTAAGACACCTGCTGTACCTCCAC 1321
QY 2658 cccacacccacccacccacccacccacccacccacccacccacccacccacccacccacccaccc 2717
DB 1322 -----CCCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1376
QY 2718 tgcgccgtttaaggcgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccg 2777
DB 1377 TCGCCCGTTTAAAGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 1436
QY 2778 taacagagggcggggaagg 2837
DB 1437 TAAACAGAGCGCGGGAACGCGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
QY 2838 cctctgacctggtgagccgaatctataaaaggaactagtccccgggcaaaaccccccgtaa 2897
DB 1497 CTTCTGCGCTCGGTGCGACCCGAATCTATAAAAGGAACCTAGTCCCGGCAAAAACCCCGTAA 1556
QY 2898 ttggagagcagagtagtggtggcgccgggagcccgagcagcagcagcagcagcagcagcagcagcagc 2957
DB 1557 TTGGAGCGAGAGTAGTGGGGCGGGGACCCCGGAGCCGAGCGGAGCCGAGCCGAGCCGAGCCGAGCC 1616
QY 2958 gctcgccgagcagggcgggcgaggctc 2984
DB 1617 GCTGGCGGAGGCGAGGGCGGGGAGGCTC 1643

RESULT 9
AC012230
LOCUS AC012230 Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
DEFINITION pieces.
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

```

SOURCE
ORGANISM
human..
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 175064)
AUTHORS
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL
Homo sapiens, clone RP11-1M10
REFERENCE
2 (bases 1 to 175064)
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galaqan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
TITLE
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 1JM10
Center clone name: 1JM10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145749 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 171264; sum-of-contigs
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1003: contig of 1003 bp in length
* 1004 1103: gap of 100 bp
* 1104 2634: contig of 1531 bp in length
* 2635 2734: gap of 100 bp
* 2735 4415: contig of 1681 bp in length
* 4416 4515: gap of 100 bp
* 4516 5785: contig of 1270 bp in length
* 5786 5885: gap of 100 bp
* 5886 7879: contig of 1994 bp in length
* 7880 7979: gap of 100 bp
* 7980 9686: contig of 1707 bp in length
* 9687 9786: gap of 100 bp
* 9787 12253: contig of 2467 bp in length
* 12254 12353: gap of 100 bp
* 12354 15228: contig of 2875 bp in length
*
* 15229 15328: gap of 100 bp
* 15329 17200: contig of 1872 bp in length
* 17201 17300: gap of 100 bp
* 17301 20131: contig of 2831 bp in length
* 20132 20231: gap of 100 bp
* 20232 22587: contig of 2356 bp in length
* 22588 22687: gap of 100 bp
* 22688 25707: contig of 3020 bp in length
* 25708 25807: gap of 100 bp
* 25808 28184: contig of 2377 bp in length
* 28185 28284: gap of 100 bp
* 28285 31338: contig of 3054 bp in length
* 31339 31438: gap of 100 bp
* 31439 34299: contig of 2861 bp in length
* 34300 34399: gap of 100 bp
* 34400 38318: contig of 3919 bp in length
* 38319 38418: gap of 100 bp
* 38419 42835: contig of 4417 bp in length
* 42836 42935: gap of 100 bp
* 42936 45448: contig of 2513 bp in length
* 45449 45548: gap of 100 bp
* 45549 48116: contig of 2568 bp in length
* 48117 48216: gap of 100 bp
* 48217 52618: contig of 4402 bp in length
* 52619 52718: gap of 100 bp
* 52719 56592: contig of 3874 bp in length
* 56593 56692: gap of 100 bp
* 56693 59635: contig of 2943 bp in length
* 59636 59735: gap of 100 bp
* 59736 63661: contig of 3926 bp in length
* 63662 63761: gap of 100 bp
* 63762 68437: contig of 4676 bp in length
* 68438 68537: gap of 100 bp
* 68538 71458: contig of 2921 bp in length
* 71459 71558: gap of 100 bp
* 71559 76888: contig of 5330 bp in length
* 76889 76988: gap of 100 bp
* 76989 82113: contig of 5125 bp in length
* 82114 82213: gap of 100 bp
* 82214 88220: contig of 6007 bp in length
* 88221 88320: gap of 100 bp
* 88321 93495: contig of 5179 bp in length
* 93500 93599: gap of 100 bp
* 93600 97901: contig of 4302 bp in length
* 97902 98001: gap of 100 bp
* 98002 103016: contig of 5015 bp in length
* 103017 103116: gap of 100 bp
* 103117 109178: contig of 6062 bp in length
* 109179 109278: gap of 100 bp
* 109279 117307: contig of 8029 bp in length
* 117308 117407: gap of 100 bp
* 117408 124079: contig of 6672 bp in length
* 124080 124179: gap of 100 bp
* 124180 131281: contig of 7102 bp in length
* 131282 131381: gap of 100 bp
* 131382 138059: contig of 6678 bp in length
* 138060 138159: gap of 100 bp
* 138160 145491: contig of 7332 bp in length
* 145492 145591: gap of 100 bp
* 145592 157391: contig of 11800 bp in length
* 157392 157491: gap of 100 bp
* 157492 175064: contig of 17573 bp in length.
*
* Location/Qualifiers
* 1. 175064
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /clone="RP11-1M10"
* /clone_lib="RPC1-11 Human Male BAC"
* /note="assembly_fragment"
* 1104. 2634
* /note="assembly_fragment"
* 2735. 4415
*
* misc_feature
* 1. 1003
* misc_feature
* 1104. 2634
* misc_feature
* 2735. 4415

```



```

/misc_feature /note="assembly_fragment"
4516. .5785
/misc_feature /note="assembly_fragment"
5886. .7879
/misc_feature /note="assembly_fragment"
7980. .9686
/misc_feature /note="assembly_fragment"
9787. .12253
/misc_feature /note="assembly_fragment"
12354. .15228
/misc_feature /note="assembly_fragment"
15329. .17200
/misc_feature /note="assembly_fragment"
17301. .20131
/misc_feature /note="assembly_fragment"
20232. .22587
/misc_feature /note="assembly_fragment"
22688. .25707
/misc_feature /note="assembly_fragment"
25808. .28184
/misc_feature /note="assembly_fragment"
28285. .31338
/misc_feature /note="assembly_fragment"
31439. .34299
/misc_feature /note="assembly_fragment"
34400. .38318
/misc_feature /note="assembly_fragment"
38419. .42835
/misc_feature /note="assembly_fragment"
42936. .45448
/misc_feature /note="assembly_fragment"
45549. .48116
/misc_feature /note="assembly_fragment"
48217. .52618
/misc_feature /note="assembly_fragment"
52719. .56592
/misc_feature /note="assembly_fragment"
56693. .59635
/misc_feature /note="assembly_fragment"
59736. .63661
/misc_feature /note="assembly_fragment"
63762. .68437
/misc_feature /note="assembly_fragment"
68538. .71458
/misc_feature /note="assembly_fragment
clone_end:SP6
vector_side:left"
71559. .76888

Query Match 38.0%; Score 1227.8; DB 2; Length 175064;
Best Local Similarity 98.7%; Pred. No. 8:2e-311;
Matches 1263; Conservative 0; Mismatches 2; Indels 14; Gaps 2;

QY 1953 agaagttggaggtctggagtggtacataattttacacgactgcaattctctgtgtgcac 2012
DB 2735 ATAAGTTGGAGTCTGGAGTGGCTACATAATTTTACACGACTGCAATTCCTGCTGCAC 2794
QY 2013 ttcacaaatgtatacaaaactaaatacaagtcgtgtgtttttatcacaggaggtgatca 2072
DB 2795 TTCACAAATGTATACAACTAAATACAAAGTCTGTTGTTTATCACAGGAGGCTGATCA 2854
QY 2073 atataatgaataaaaggggctgtccatattgttctgtgtgtgtgtgtgtgtgtgtttt 2132
DB 2855 ATATAAGTAATTAAGAGGGGCTGGTCCATATTTGTTGTTGTTTGTGTTGTTTGTG 2905
QY 2133 gttgtttcttttttttttttttttttttttttttttttttttttttttttttttttttt 2192
DB 2906 GTTTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 2965
QY 2193 agatgttctctctgggtctctgaggaacctgaggaagctgggagctcaggctgggaatctccaaagc 2252
DB 2966 AGATGTTCCTCTCGGGTCTCTGAGGACCTGAGGAGCTGGGGAGTCTGAGGAGTCTGAGG 3025

```

```

QY 2253 agtagtgcgctatcaaaaatcaaaatccaggtttgtggggggaacaaacacagccca 2312
DB 3026 AGTAGGTGCGCTATCAAAAATCAAAAGTCCAGGTTTGTGGGGGGAACAAACAGACGCCA 3085
QY 2313 ttacccagaggaactgtccgcttcctccctccctccctccctccctccctccctccctcc 2372
DB 3086 TTACCCAGAGGAGTGTCCGCTTCCCTCACCACAGCCTAGGCTTGTGAAGGAAACAAA 3145
QY 2373 agacaagacaaaatgattggcgtccctgaggaagatcagcctagagctctctctccccc 2432
DB 3146 AGACAAGACAAAATGATTGGCGTCTGAGGAGATTACGCTTAGAGCTCTCTCTCCCCCA 3205
QY 2433 atcctccctccggtgaggaatacaaaagaaataaaatgaggaagcagattt 2492
DB 3206 ATCCCTCCCTCCGCTGAGGAATAACAAAGGAAATAATTTGGGAAGCAGGATTT 3265
QY 2493 agaggaaagcaaatccacttggtgcttccctggctcccggaacgtggactagagagtcgtg 2552
DB 3266 AGAGGAAGCAAAATTCACACTGCTGCTTGGCTGCCGGAACGTGGACTAGAGAGTCTGCG 3325
QY 2553 gcgagcccgagcccgagcgttcccgcggttcttaggcggcgccggcgccggcgccgggaa 2612
DB 3326 GCGAGCCCGGAGCCGAGCGCTTCCGCGCGCTTTAGGCGCGGCGCCGCGCGGGGAA 3385
QY 2613 gggagcgagacgcgagaccctaaagacacctgctgtacctccaccccccaccccccaccc 2672
DB 3386 GGGAGCGCAGACCGCGGACCTAAGACACCTGCTGTACCTCCAC ----CCCCACCCCA 3440
QY 2673 cccactccccccaaactccctagatgtgctgtgggggctgaaactgcgcgttttaagg 2732
DB 3441 CCCACCTCCCCCAACTCCCTAGATGTGCTGGGGGCTGAACGTCGCCCTTTAAGGG 3500
QY 2733 gcgggcccggtccacgctgtcttctgctgagtgactgaactacataaagagggcg 2792
DB 3501 GCGGGCCCGGCTCCACGCTTTTCTGCTGAGTGAACCTATATAAACAGAGCGGG 3560
QY 2793 aagggggcgaggagagagacagcctttgacgtagtaacatctcgcctcggtg 2852
DB 3561 AACGGGCGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTG 3620
QY 2853 cagccgaactataaaggaaactagtcctcgcgcaaaaccccgtaattgagcgagagtg 2912
DB 3621 CAGCCGAATCTATAAAGGAACCTAGTCCCGGCAAAACCCCGTAATTGCGAGCGAGGTG 3680
QY 2913 agtggggcgagcccgagagccgagccctctctctctctctctctctctctctctctctct 2972
DB 3681 AGTGGGCGCGGACCCGCGAGCGCAGCCCTTCTCTCCCGGCTGCGGCGAGGCGAG 3740
QY 2973 ggcgggagctcgcgacacacagacggcttctcagggcgcttgcctctgtttttt 3032
DB 3741 GCGGGGAGCTCCGCGCACCAACAGAGCGGTTCTCAGGGCGCTTTGCTCTTGTGTTT 3800
QY 3033 ccccggttctgttttctcccttctcggaggtctgtcaagggttaggagaaagagacg 3092
DB 3801 CCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAAGGGGTAGGAGAAAGAGCG 3860
QY 3093 caaacacaaagtggaaaacaggttaagagctctccagtgacttaacttggggttattgt 3152
DB 3861 CAAACAAAAGTGAACACAGGTAAGAGGCTCTCCAGTGACTTACTTGGCGCTTATTGT 3920
QY 3153 ttgtttcagggcgaagggcttcgggaagtgcgttcggttcgggacttgcagcgag 3212
DB 3921 TTTGTTTCAGGGCCAGGAGGCTTCGGGAAGTGTCTCGGTTTCGGGACTTTGATCCGAG 3980
QY 3213 cccacatccccaccactt 3231
DB 3981 CCCCACATCCCCACCACCTT 3999

```

RESULT 10
HSA252201
LOCUS
DEFINITION

Homo sapiens partial ABC-1 gene for ATP-binding cassette
1167 bp DNA linear PRI 10-APR-2001

JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
 REFERENCE 2 (bases 224 to 1167)
 AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
 Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
 TITLE Direct Submission
 JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 REFERENCE 3 (bases 1 to 1167)
 AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
 Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 REMARK Sequence update by submitter
 COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.
 FEATURES
 Location/Qualifiers
 1..1167
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q31"
 /224..844
 /gene="ABCA1"
 845..1147
 /number=1
 BASE COUNT 278 a 313 c 328 g 248 t
 ORIGIN
 Query Match 35.0%; Score 1130.4; DB 9; Length 1167;
 Best Local Similarity 98.7%; Pred. No. 9.3e-286;
 Matches 1165; Conservative 0; Mismatches 1; Indels 14; Gaps 2;
 QY 1955 aagttggaggtctggagtggtacataattttacacgactgcaattctctggtgcactt 2014
 Db 2 AAGTTGGAGGCTGGAGTGGCTACATAATTTTACACGACTGCAATTTCTCTGGCTGCAC 61
 QY 2015 cacaaatatacaactaaatacaagtcctgtgtttttatcacagggaggtgatacaat 2074
 Db 62 CACAAATGTATACAAACTAAATACAAAGTCCCTGTGTTTATCACAGGGAGGCTGATCA 121
 QY 2075 ataatgaattaaaaggggctgtcccatattgtctgtgtttgtttgtttgtttgtt 2134
 Db 122 ATAAATGAATTAAGGGGGCTGGTCCATATTTGTCGTGTTTTG-----TTTGT 172
 QY 2135 ttgtttctttttgtttttgtttgttgcctctctctcatttataagagaagcagtaag 2194
 Db 173 TTGTTTCTTTTGTGTTTGTGGCTCCTTCTCTCAATTTATGAAGAGAGCAGTAAG 232
 QY 2195 atgttctctcgggtcctctgaggacctgggagctcaggctggaatctccaagcag 2254
 Db 233 ATGTTCTCTCGGTCCTCTGAGGACCTGGGAGCTCAGGCTGGGATCTCCAAGGCAG 292
 QY 2255 taggtcgactcaaaaatcaaaagtcagggtttgtgggggaaacaaaagcagccatt 2314
 Db 293 TAGGTCGCTATCAAAATCAAAATCCAGGTTTGTGGGGGAAACAAACACAGCCCAT 352
 QY 2315 acccagagactgtccgcttctcctccctccctccctccctccctccctccctccct 2374
 Db 353 ACCCAGAGACTGTCCGCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 412
 QY 2375 acaagacaaaatgattggctcctgaggagattcagcctagactctctctccccaat 2434
 Db 413 ACAAGACAAAATGATTGGCGCTCTGAGGAGATTACAGCTAGACTCTCTCTCCCAAT 472
 QY 2435 cctcctcctcgggtcctgagaaactcaaaaggaagaaataattcggaagcaggtattag 2494
 Db 473 CCTCTCTCTCGGCTGAGGAACCTACAAAGGAAAGAAATTCGGAAGCAGGATTAG 532
 QY 2495 aggaagcaaatccactggtgcttggctggtccgggaacgtggaactagagagctcgccg 2554
 |||||||

Db 533 AGGAAGCAAAATCCACTGGTGCCTTTGGCTGCGGGAACGCTGGACTAGAGAGTCTCGGC 592
 QY 2555 gcagccccagagccagcgttcccgcgctcttagccgagccgagccgagggggaag 2614
 |||||||
 Db 593 GCAGCCCGAGCCAGCGGCTTCCCGCGCTTTAGCCGCGCGCGCGGGGGGAGG 652
 QY 2615 ggagcagagccgagcccttaagacacacctgtctacccctccacccccccccacc 2674
 |||||||
 Db 653 GGAGCGACAGCCGCGACCCCTAAGACACCTGCTGTACCTCCAC-----CCCCACCC 707
 QY 2675 caactcccccaactccctagatgtgtggtggtggtggtggtggtggtggtggtggt 2734
 |||||||
 Db 708 CACTTCCCCCAACTCCCTAGATGTGTGCGCGCGCTGAACGCTGCCCGTTTAAAGGG 767
 QY 2735 ggccccgggtccacgtgcttctgctgagtgactgaactacataaacagagccggaa 2794
 |||||||
 Db 768 GGGCCCGGCTCCACGTGCTTTCTGTAGTGACTGAACCTACATAAACAGAGCCCGGAA 827
 QY 2795 ggggggggag 2854
 |||||||
 Db 828 GGGGGCGGAGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGGCTCGGTGCA 887
 QY 2855 gccgaatctataaaaggaactagtcctcgcaaaaccccccaactgagagagagagag 2914
 |||||||
 Db 888 GCGGAATCTATAAAAGGAACTAGTCTCGGCAAAACCCCGTAATTCGAGCGAGAGTGAG 947
 QY 2915 tggggcgagccgag 2974
 |||||||
 Db 948 TGGGGCGGAGCCCGAGAGCCGAGCCCTTCTCTCCGGGCTGCGCAGGAGGAGG 1007
 QY 2975 cggggagctcggcgacacacagagcgggttctcagggcgcttctgcttcttcttctt 3034
 |||||||
 Db 1008 CGGGAGCTCCGCGACCAACAGACGCGGTCTCAGGCGCTTTGCTCTGTTTTC 1067
 QY 3035 ccggttctgtttctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 3094
 |||||||
 Db 1068 CCGGTTCTGTTTCTCCTCCCTTCTCCGGAAGGCTTTCAGGGGTAGGAGAAAGAGACGCA 1127
 QY 3095 aacacaaaagtaaaacagagtaagagctctccagtgac 3134
 |||||||
 Db 1128 AACAAAAAGTGAAGAACAGTAAGAGGCTCTCCAGTGAC 1167

RESULT 12

AC021246 69570 bp DNA linear HTG 13-JUL-2000
 LOCUS Homo sapiens clone RP11-IN10, LOW-PASS SEQUENCE SAMPLING.
 AC021246
 DEFINITION Homo sapiens chromosome, clone RP11-IN10
 AC021246
 ACCESSION
 VERSION AC021246.2 GI:9119882
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 69570)

AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens chromosome, clone RP11-IN10

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 69570)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Becker, R., Bada, F.,
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
 Choe, Y., Collangelo, M., Collins, S., Collins, S., Collymore, A., Cooke, P.,
 DeAtellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
 Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
 Garfield, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lehotsky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,

Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6705871.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N_10

* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 872 971: gap of 100 bp in length
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length
* 3746 3845: gap of 100 bp
* 3846 4696: contig of 851 bp in length
* 4697 4796: gap of 100 bp
* 4797 5640: contig of 844 bp in length
* 5641 5740: gap of 100 bp
* 5741 6540: contig of 800 bp in length
* 6541 6640: gap of 100 bp
* 6641 7509: contig of 869 bp in length
* 7510 7609: gap of 100 bp
* 7610 8479: contig of 870 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 880 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length
* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length

* 20922 21021: gap of 100 bp
* 21022 21865: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 26621: contig of 851 bp in length
* 26622 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 27676: gap of 100 bp
* 27677 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 861 bp in length
* 55230 55329: gap of 100 bp

Best Local Similarity 84.08; Pred. No. 5.7e-134;
Matches 569; Conservative 0; Mismatches 107; Indels 1;

Qy 2 caggagcattggtgagcagctgctgaatctcaattactcaggaggtgaggtgcaatgag 61
Db 1678 CCGGCGATGGTGGCAGGTGCTGTAATCTCAGCTACTCGGAGGTGGAGGTGCAATGAG 1619
Qy 62 ccagatcgaccattgacatccagctggcgcaaaagtgaaactccatctcaatta 121
Db 1618 CCCAGATCGCACCATTGACATCCAGCTGGCGCAAAAGTGAACCTCCATCTCANTTA 1559
Qy 122 aaaaaaagaatgatttgggtgctgaactcgaataggttaggagaagaaggagagga 181
Db 1558 AAAAAAATAATGATTGTTGGTGTGCTGCACTTCAATAGTAGGAGAAGAAGAGAGAGGA 1499
Qy 182 gatgagaggtcaggagatcaattactctctaaatcatctaggaagaaataacacctt 241
Db 1498 GATGAGAGGTGAGGAGATCAATTAATCTCTAAATCATCTAGGAAAGATAACACCTT 1439
Qy 242 ttaataacactctctgttttataacatcattctgccaagagctcaaaagtgttcaaca 301
Db 1438 TTAATAACACTCTCTGCTTTTATAACATCATCTGCCAAGGAGCTCAAGGTTTCAACAA 1379
Qy 302 agttcaattcagaaaaaccccttgaggagaagaataatatacatcttctccattttaa 361
Db 1378 AGTTCACTTTTCAGAAAAACCCCTTTTGGAGGAGACAGAAATATACATCTTCTCCATTTTAA 1319
Qy 362 agatgaagaacagcgagcagcaatagctgaatcctgtaatccagacactttggagggc 421
Db 1318 AGATGAAGAACAGCGCGGCGCAATAGGCTAATGCTGTAAATCCAGCAGCTTTGGGAGGC 1259
Qy 422 tgagccagagatcgcttgagctccagagttgagaccagcctgggataacatggcaaaa 481
Db 1258 TGAGGCCAAAGGATCGCTTGAGCTCCAGAGTTTGAGACCAGCTTGGATAACATGCAAAA 1199
Qy 482 cctgtctctcaaaaaataacaaaaattagatgggtggtgagcagcctgtggtc 541
Db 1198 CCCTGCTCTCAAAAAAATAGGAAAAATAGATGGGTGTTGGTGGCATGCACTGTGTC 1139
Qy 542 ccagctacttggagcctgaagtgaggatcgcttgagccagggagtgcaagtctacac 601
Db 1138 CCAGCTAC-TGGGAGCTTGAAGTGGGAGGATCGCTTTNNNNNNNNNNNNNNNNNNNN 1080
Qy 602 tgagccatgattgactcactcactcagcctggtgtagacagagcaagaccctgtctca 661
Db 1079 NNN 1020
Qy 662 aaaaaaagaatgaaaga 678
Db 1019 NNNNNNNNNNNNNNNNA 1003

RESULT 14
AC021246/c
LOCUS AC021246 69570 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC021246
VERSION AC021246.2 GI:9119882
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69570)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,I., Boukhvalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,I., Boukhvalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,

Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Maconald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE JOURNAL

COMMENT

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N10

* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
872 971: gap of 871 bp in length
872 971: gap of 100 bp
972 1834: contig of 863 bp in length
1835 1934: gap of 100 bp
1935 2804: contig of 870 bp in length
2805 2904: gap of 100 bp
2905 3745: contig of 841 bp in length
3746 3845: gap of 100 bp
3846 4696: contig of 851 bp in length
4697 4796: gap of 100 bp
4797 5640: contig of 844 bp in length
5641 5740: gap of 100 bp
5741 6540: contig of 800 bp in length
6541 6640: gap of 100 bp
6641 7509: contig of 869 bp in length
7510 7609: gap of 100 bp
7610 8479: contig of 870 bp in length
8480 8579: gap of 100 bp
8580 9430: contig of 851 bp in length
9431 9530: gap of 100 bp
9531 10376: contig of 846 bp in length
10377 10476: gap of 100 bp
10477 11322: contig of 846 bp in length
11323 11422: gap of 100 bp
11423 12302: contig of 880 bp in length
12303 12402: gap of 100 bp
12403 13280: contig of 878 bp in length
13281 13380: gap of 100 bp
13381 14241: contig of 861 bp in length
14242 14341: gap of 100 bp
14342 15196: contig of 855 bp in length
15197 15296: gap of 100 bp
15297 16123: contig of 827 bp in length
16124 16223: gap of 100 bp


```

Db 67162 GCGATAGTGTGTTTCGAGGCGCAAGAGGCTTCGAAAGTGTCTGGTTCGGGACTT 67103
QY 3203 tgatccgagaccacatcccc 3224
Db 67102 TGATCCGAGGCCACATCCCC 67081

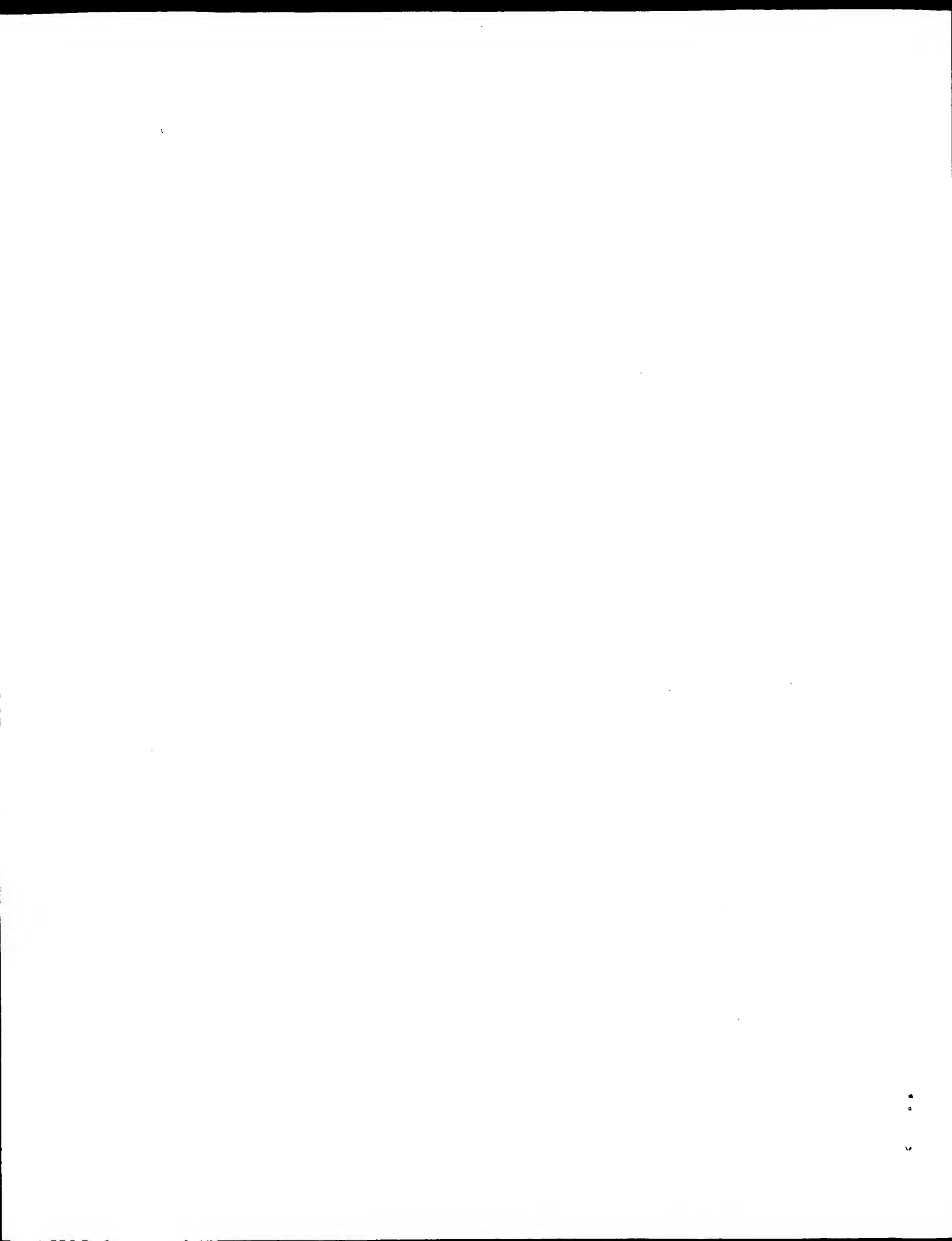
RESULT 15
AC021345
LOCUS Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC021345
ACCESSION AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 90698)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-24J9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 90698)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeAtellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrel,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lander,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,I., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705761.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4483
Center clone name: 24_J_9
-----
* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1073: contig of 863 bp in length
* 1074 1073: gap of 100 bp
* 1074 2824: contig of 851 bp in length

```

```

2825 2924: gap of 100 bp
2925 3802: contig of 878 bp in length
3803 3902: gap of 100 bp
3903 4816: contig of 914 bp in length
4817 4916: gap of 100 bp
4917 5759: contig of 843 bp in length
5760 5859: gap of 100 bp
5860 6764: contig of 905 bp in length
6765 6864: gap of 100 bp
6865 7747: contig of 883 bp in length
7748 7847: gap of 100 bp
7848 8755: contig of 908 bp in length
8756 8855: gap of 100 bp
8856 9753: contig of 898 bp in length
9754 9853: gap of 100 bp
9854 10757: contig of 904 bp in length
10758 10857: gap of 100 bp
10858 11732: contig of 875 bp in length
11733 11832: gap of 100 bp
11833 12739: contig of 907 bp in length
12740 12839: gap of 100 bp
12840 13710: contig of 871 bp in length
13711 13810: gap of 100 bp
13811 14684: contig of 874 bp in length
14685 14784: gap of 100 bp
14785 15662: contig of 878 bp in length
15663 15762: gap of 100 bp
15763 16677: contig of 915 bp in length
16678 16777: gap of 100 bp
16778 17678: contig of 901 bp in length
17679 17778: gap of 100 bp
17779 18679: contig of 901 bp in length
18680 18779: gap of 100 bp
18780 19632: contig of 853 bp in length
19633 19732: gap of 100 bp
19733 20634: contig of 902 bp in length
20635 20734: gap of 100 bp
20735 21620: contig of 886 bp in length
21621 21720: gap of 100 bp
21721 22579: contig of 859 bp in length
22580 22679: gap of 100 bp
22680 23568: contig of 889 bp in length
23569 23668: gap of 100 bp
23669 24554: contig of 886 bp in length
24555 24654: gap of 100 bp
24655 25521: contig of 867 bp in length
25522 25621: gap of 100 bp
25622 26487: contig of 866 bp in length
26488 26587: gap of 100 bp
26588 27464: contig of 877 bp in length
27465 27564: gap of 100 bp
27565 28466: contig of 902 bp in length
28467 28566: gap of 100 bp
28567 29464: contig of 898 bp in length
29465 29564: gap of 100 bp
29565 30447: contig of 883 bp in length
30448 30547: gap of 100 bp
30548 31453: contig of 906 bp in length
31454 31553: gap of 100 bp
31554 32452: contig of 899 bp in length
32453 32552: gap of 100 bp
32553 33447: contig of 895 bp in length
33448 33547: gap of 100 bp
33548 34435: contig of 888 bp in length
34436 34535: gap of 100 bp
34536 35433: contig of 898 bp in length
35434 35533: gap of 100 bp
35534 36440: contig of 907 bp in length
36441 36540: gap of 100 bp
36541 37423: contig of 882 bp in length
37423 37522: gap of 100 bp
37523 38402: contig of 880 bp in length
38403 38502: gap of 100 bp

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 00:12:11 ; Search time 520.94 Seconds
(without alignments)
10648.732 Million cell updates/sec

Title: US-09-846-456-1
Perfect score: 3231
Sequence: 1 acaggcatgtgcagggtg.....gccccacatccccacactt 3231

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3127.8	96.8	183999	22 AAF92831	Human ABC1 genomic
2	1485.4	46.0	1643	22 AAF24681	Nucleotide sequenc
3	1485.4	46.0	1643	22 AAF24703	Nucleotide sequenc
4	336	10.4	763	22 AAF04729	Human cDNA clone (
5	336	10.4	1750	22 AAF17451	Human cDNA sequenc
6	330.8	10.2	20645	22 AAF05355	Human reproductive
7	330.6	10.2	5076	22 AAF16699	Human nervous syst
8	325.8	10.1	5075	22 AAF16701	Human nervous syst
9	323.8	10.0	36785	22 AAF82208	Human immune/haema

c 10	320.4	9.9	23934	22	ABAL9145	Human nervous syst
c 11	320.4	9.9	23934	22	AAAL36171	Human musculoskele
c 12	320.4	9.9	23934	22	AAAL36179	Human musculoskele
c 13	320.4	9.9	23934	22	AAAL04522	Human reproductive
c 14	320.4	9.9	23934	22	AAAS28343	Genomic sequence #
c 15	320.4	9.9	23934	22	AAK71442	Human immune/haema
c 16	313.6	9.7	26390	22	AAK65971	Human immune/haema
c 17	312.8	9.7	9745	22	AAAS6759	Human cardiovascular
c 18	312.8	9.7	12149	22	AAAS6758	Human cardiovascular
c 19	312.4	9.7	8663	22	AAAS0133	Human lung antigen
c 20	312.2	9.7	15275	22	AAAS5975	Human cardiovascular
c 21	311.8	9.7	9731	22	AAK72933	Human immune/haema
c 22	311.8	9.7	9731	22	AAK85096	Human immune/haema
c 23	311.8	9.7	9733	22	AAK72935	Human immune/haema
c 24	311.8	9.7	9733	22	AAK85097	Human immune/haema
c 25	311.6	9.6	16555	22	AAK70102	Human immune/haema
c 26	311.6	9.6	16555	22	AAK73172	Human immune/haema
c 27	310.4	9.6	5797	22	ABAL5723	Human nervous syst
c 28	310.4	9.6	5797	22	AAAL03405	Human reproductive
c 29	309.8	9.6	23885	22	AAK70103	Human immune/haema
c 30	309.8	9.6	23885	22	AAK73173	Human immune/haema
c 31	309.8	9.6	66933	22	ABAB2625	Human HBM gene reg
c 32	309.8	9.6	72049	22	ABAB2623	Human HBM gene reg
c 33	307.2	9.5	15555	22	AAK73539	Human immune/haema
c 34	307.2	9.5	15558	22	AAK73538	Human immune/haema
c 35	306.2	9.5	23241	22	AAK97870	Human neuroblastom
c 36	306.2	9.5	23241	22	AAK97871	Human neuroblastom
c 37	304.8	9.4	15554	22	AAK73537	Human immune/haema
c 38	304.6	9.4	2743	22	AAK86888	Human immune/haema
c 39	304.2	9.4	8948	22	AAK67209	Human immune/haema
c 40	304	9.4	45000	22	AAAL2437	DNA encoding 1-ami
c 41	303.8	9.4	32191	22	AAAS0497	DNA encoding novel
c 42	303.8	9.4	32191	22	AAAL06277	Human reproductive
c 43	303.8	9.4	46107	22	AAK71730	Human immune/haema
c 44	303.4	9.4	32224	22	AAK89986	Human digestive sy
c 45	303.2	9.4	26372	22	AAK77103	Human immune/haema

ALIGNMENTS

RESULT 1	
AAAF92831	
ID AAF92831 standard; DNA; 183999 BP.	
XX AAF92831;	
AC AAF92831;	
XX	
DT 17-MAY-2001 (first entry)	
XX	
DE Human ABC1 genomic DNA.	
XX	
KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.	
XX	
OS Homo sapiens.	
XX	
PN WO200115676-A2.	
XX	
PD 08-MAR-2001.	
XX	
PF 01-SEP-2000; 2000WO-IB01492.	
XX	
PR 01-SEP-1999; 99US-0151977.	
PR 15-MAR-2000; 2000US-0526193.	
PR 23-JUN-2000; 2000US-0213958.	
XX	
(UYBR-) UNIV BRITISH COLUMBIA.	
(XENO-) XENON GENETICS INC.	
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;	
XX WPI; 2001-244356/25.	
PT Treating a lower than normal high density lipoprotein-cholesterol	

PT (HDL-C) level, a higher than normal triglyceride level, or a
 PT cardiovascular disease, by administering a compound that modulates LXR-
 XX or RXR-mediated transcriptional activity -
 PS Claim 8; Fig 1; 317pp; English.

XX The present invention relates to a method for treating a patient
 CC diagnosed as having a lower than normal high density
 CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
 CC triglyceride level, or a cardiovascular disease, involving
 CC administering a compound that modulates LXR- or RXR-mediated
 CC transcriptional activity or ABC1 expression or activity.
 CC The LXR gene product may be used in an assay to identify
 CC compounds useful for the treatment of a disease or condition selected a
 CC lower than normal HDL cholesterol level, a higher than normal
 CC triglyceride level, and a cardiovascular disease.

XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 96.8%; Score 3127.8; DB 22; Length 183999;
 Best Local Similarity 98.0%; Pred. No. 0;
 Matches 3183; Conservative 10; Mismatches 37; Indels 18; Gaps 3;

QY 2 caggagcagtggtgagtgctgtaactcagttactcagttactcggagtgagtggaatgag 61
 DB 25880 ccggcagtggtgagtgctgtaactcagttactcagttactcggagtgagtggaatgag 25939
 QY 62 ccagagtcgacattgacatcagctcggcgaacaaagggtgaactcactcaatta 121
 DB 25940 ccagagtcgacattgacatcagctcggcgaacaaagggtgaactcactcaatta 25999
 QY 122 aaaaaaagaatgatttgggtgagcttcaaataggtaggaggaagagagagga 181
 DB 26000 aaaaaaagaatgatttgggtgagcttcaaataggtaggaggaagagagagga 26059
 QY 182 gatggagggcagggagatgataactatcctaaatcagtaggaggaagataacacct 241
 DB 26060 gatggagggcagggagatgataactatcctaaatcagtaggaggaagataacacct 26119
 QY 242 ttaatacactctctgttttataacatcattctgcaaggagctcaaaagggttcaaca 301
 DB 26120 ttaatacactctctgttttataacatcattctgcaaggagctcaaaagggttcaaca 26179
 QY 302 agttcatttcaaaaaaccccttgaggagacagataatatactctctctcatttaa 361
 DB 26180 agttcatttcaaaaaaccccttgaggagacagataatatactctctcatttaa 26239
 QY 362 agatgaagaacagggcgacaaatggttaactgctgaatccagcactttgggagc 421
 DB 26240 aga tgaagaacagggcgacaaatggttaactgctgaatccagcactttgggagc 26299
 QY 422 tgaggccagagatcgcttgagctccagagtttgagaccagctggatgaacatggcaaa 481
 DB 26300 tgaggccasargatcgcttgagctccaragtttgagaccagctggatgaacatggcaaa 26359
 QY 482 cctgtctctacaaaaatacaaaaattagatgggtggtggcagcacttggtg 541
 DB 26360 cctgtctctacaaaaatacaaaaattagatgggtggtggcagcacttggtg 26419
 QY 542 ccagcacttgaggagtgtaagggtggagagatcgcttgagccagagagtaagttacac 601
 DB 26420 ccagcacttgaggagtgtaagggtggagagatcgcttgagccagagagtaagttacac 26479
 QY 602 tgagccatgattgagcactgcactccagcctgggttagacagagcagacccctgtctaa 661
 DB 26480 tgagccatgattgagcactgcactccagcctgggttagacagagcagacccctgtctaa 26539
 QY 662 aaaaaaagaatgaaagaaagaaagaa -----agaggagagagagagatga 710
 DB 26540 acaaaa tgaatgaaagaaagaaagaaagagagtgagagagagagatgagggagggga 26599
 QY 711 ggggagaggggggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 770

DB 26600 gggtgagcaggggggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 26659
 QY 771 aaaaaaacaagatgaacagagcagaaagactttacgttaaatctctcatcatggtggtg 830
 DB 26660 aaatacgaacaatgaacagagcagaaagactttacgttaaatctctcatcatggtggtg 26719
 QY 831 tcaagtttgaccccccaaaaccccaattattgaccaaggttattctttgactgagggcaagg 890
 DB 26720 tcaagtttgaccccccaaaaccccaattattgaccaaggttattctttgactgagggcaagg 26779
 QY 891 ggtcagctctcctgggcttggtttgagaaagctcactctggtcctttctgagatccat 950
 DB 26780 ggtcagctctcctgggcttggtttgagaaagctcactctggtcctttctgagatccat 26839
 QY 951 cctttctttttattttcttgacacgagcttctctctgctcactcaggtgagtgagcag 1010
 DB 26840 cctttctttttattttcttgacacgagcttctctctgctcactcaggtgagtgagcag 26899
 QY 1011 tggcatgatctcgactcactgtaacctgctcccggtgtcaagcgattctctcctgcctc 1070
 DB 26900 tggcatgatctcgactcactgtaacctgctcccggtgtcaagcgattctctcctgcctc 26959
 QY 1071 agcctcctgagataacagggcgcccccacacacatctggtcctaatttttttagtaa 1130
 DB 26960 agcctcctgagataacagggcgctccacacacatctggtcctaatttttttagtaa 27019
 QY 1131 agactgggtttctcatcatgttggccaggttgggttcgaactcctgacctgaggtgagctgc 1190
 DB 27020 agactgggtttctcatcatgttggccaggttgggttcgaactcctgacctgaggtgagctgc 27079
 QY 1191 ccactttggctcccaaaagtgctgggattacagggcatgagccactgcgccagctcagat 1250
 DB 27080 ccactttggctcccaaaagtgctgggattacagggcatgagccactgcgccagctcagat 27139
 QY 1251 ccactcctttcaggggcaaaacagtcctcctggtgcaaaagggccatgcccacagagttat 1310
 DB 27140 ccactcctttcaggggcaaaacagtcctcctggtgcaaaagggccatgcccacagagttat 27199
 QY 1311 ggtacctgggactccagaattccttgcctggtgctccacacatgaccttccagggcctg 1370
 DB 27200 ggtacctgggactccagaattccttgcctggtgctccacacatgaccttccagggcctg 27259
 QY 1371 cttgggctctctctatgctgctgctgagtggtgtagaaccactgagtgagtaactg 1430
 DB 27260 cttgggctctctctatgctgctgctgagtggtgtagaaccactgagtgagtaactg 27319
 QY 1431 ggcttgagccgctgctgagatcctgctgactgtagcatgaggggcttggcagctg 1490
 DB 27320 ggcttgagccgctgctgagatcctgctgactgtagcatgaggggcttggcagctg 27379
 QY 1491 aatgtctgcatcaggtggtgggagttcttgaataatgagctggagctgggaggtgggaagaga 1550
 DB 27380 aatgtctgagtgaggtggtgggagttcttgaataatgagctggagctgggaggtgggaagaga 27439
 QY 1551 agtaggcttggggagctctctcactgcacacatcttgcgcaaaactcaggtcaaaactg 1610
 DB 27440 agtaggcttggggagctctctcactgcacacatcttgcgcaaaactcaggtcaaaactg 27499
 QY 1611 tgaagagctcaaatgtaactgctccctcaaggtggtgctacaaaggtatcttctcaaggt 1670
 DB 27500 tgaagagctcaaatgtaactgctccctcaaggtggtgctacaaaggtatcttctcaaggt 27559
 QY 1671 agtaggacttgggctccactgacgtgacgttccagggcctggtggcctctctctacaggtc 1730
 DB 27560 agtaggacttgggctccactgacgtgacgttccagggcctggtggcctctctctacaggtc 27619
 QY 1731 tgcctgagcttctctatgaa -----tccctcagggcagagatcatttagactcttca 1785
 DB 27620 tgcctgagcttctctatgaa -----tccctcagggcagagatcatttagactcttca 27679
 QY 1786 gtttagctgaggttggccagaaataaaggtgacatttagttgttggctgagtgatgac 1845

Db 27680 gtttgacgtgagttttgtccagaataaagtgacatttagttgttgcttgatgac 27739
QY 1846 ttaaatatttagac--atgtgtgttagcctgcattctactctctgtcttttttttggcc 1903
Db 27740 ttaaatatttagacataatgtgtgttagcctgcattctactctctgtcttttttttggcc 27799
QY 1904 cctccagtggtttgtgtagttttgtctccctacagcgaagcaaaacagagaagtggag 1963
Db 27800 cctccagtggtttgtgtagttttgtctccctacagcgaagcaaaacagagaagtggag 27859
QY 1964 gcttgagtggtacataattttacagagactgcaattctctggtgcaattcacaaatgt 2023
Db 27860 gcttgagtggtacataattttacagagactgcaattctctggtgcaattcacaaatgt 27919
QY 2024 atacaactaaatacaagctgctgtttttatcacagagagctgatacataataatgaaa 2083
Db 27920 atacaactaaatacaagctgctgtttttatcacagagagctgatacataataatgaaa 27979
QY 2084 ttaaaaggggctgctcattattgttctgtgtttttgtttttgtttttgttttttctt 2143
Db 27980 ttaaaaggggctgctcattattgttctgtgtttttgtttttgtttttttttnnn 28039
QY 2144 tttttgtttttgtgctcctctctcctcaatttatgaagaagacagtaagatttctc 2203
Db 28040 tnnntgtttttgtgctcctctctcctcaatttatgaagaagacagtaagatttctc 28099
QY 2204 tcgggtcctctgagggacacctgggagctcaggctgggaatctccaagcagtagtgctc 2263
Db 28100 tcgggtcctctgagggacacctgggagctcaggctgggaatctccaagcagtagtgctc 28159
QY 2264 tatcaaaatcaaatccaggtttgtgtgggggaaacaaaagcagccattaccacaggg 2323
Db 28160 tatcaaaatcaaatccaggtttgtgtgggggaaacaaaagcagccattaccacaggg 28219
QY 2324 actgtccgcttccctcaaccagctaggcctttgaaagaaacaaaagacagaacaa 2383
Db 28220 actgtccgcttccctcaaccagctaggcctttgaaagaaacaaaagacagaacaa 28279
QY 2384 aatgattggcctcctgagggagattcagctagagctctctccccaatccctccctc 2443
Db 28280 aatgattggcctcctgagggagattcagctagagctctctccccaatccctccctc 28339
QY 2444 cggctgaggaactaacaagaaagaaataatttcggaagcagagatttagagaagcaa 2503
Db 28340 cggctgaggaactaacaagaaagaaataatttcggaagcagagatttagagaagcaa 28399
QY 2504 attccactggtccttgctgctgctgctgctgctgctgctgctgctgctgctgctg 2563
Db 28400 attccactggtccttgctgctgctgctgctgctgctgctgctgctgctgctgctg 28459
QY 2564 agccagcgtctcccgctgctttaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2623
Db 28460 agccagcgtctcccgctgctttaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 28519
QY 2624 ccgcggaacctaaagacacctgtgttaccctccaccccccaccccccaccccccaccc 2683
Db 28520 ccgcggaacctaaagacacctgtgttaccctccaccccccaccccccaccccccaccc 28579
QY 2684 ccaactccctagatgt 2743
Db 28580 ccaactccctagatgt 28639
QY 2744 ctccagctgtcttctgctgagtgactgaactacataaagagccgcgcgcgcgcgcgcgc 2803
Db 28640 ctccagctgtcttctgctgagtgactgaactacataaagagccgcgcgcgcgcgcgcgc 28699
QY 2804 gaggaggagagacacagctgtttgaccgtagtagtaacctctgcgctggtgcagcgcaatct 2863
Db 28700 gaggaggagagacacagctgtttgaccgtagtagtaacctctgcgctggtgcagcgcaatct 28759
QY 2864 ataaaggaactagtcctccggcaaaaccccgtaattgcgagcagagtgagtgggccgg 2923
Db 28760 ataaaggaactagtcctccggcaaaaccccgtaattgcgagcagagtgagtgggccgg 28819

QY 2924 gacccagagccgagccgaccccttctcccggtgtgcgagcagcagcagcagcagcagc 2983
Db 28820 gacccagagccgagccgaccccttctcccggtgtgcgagcagcagcagcagcagc 28879
QY 2984 ccgcgacccaacagagcgggtctctcagggcgtttgctccttcttttcccggttctg 3043
Db 28880 ccgcgacccaacagagcgggtctctcagggcgtttgctccttcttttcccggttctg 28939
QY 3044 ttcttcccttctccggaagcctgtcgaagggttaggagaaagagacgcaaacacaaa 3103
Db 28940 ttcttcccttctccggaagcctgtcgaagggttaggagaaagagacgcamacacaaa 28999
QY 3104 gtgaaacacaggttaagagcgtctccagtgacttacttggcgttattgtttgttcgag 3163
Db 29000 gtgaaacacaggttaagagcgtctccagtgacttacttggcgttattgtttgttcgag 29059
QY 3164 gcaagagcgtctccggaagcgtcgtcgttccgggactttgatcccgagccacacccc 3223
Db 29060 gcaagagcgtctccggaagcgtcgtcgttccgggactttgatcccgagccacacccc 29119
QY 3224 caccactt 3231
Db 29120 caccactt 29127

RESULT 2

AAF24681

ID AAF24681 standard; DNA; 1643 BP.

XX AAF24681;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

Human; adenosine triphosphate binding cassette protein 1; ABC1;
XW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX WO200078972-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/14.

Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
useful for the development of agents for the treatment of heart disease
and other disorders associated with hypercholesterolemia and
atherosclerosis -

PS Claim 1; Page 143-144; 215pp; English.

XX The present sequence represents the 5' flanking region of the human
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
CC variety of substrates across the plasma membrane. ABC1 is a pivotal
CC protein in the apolipoprotein-mediated mobilisation of intracellular
CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1

QY 2838 cctctgcgtcgtcagccgaactataaaaggaactagtcctccggcaaaaaccccgtaa 2897
 |||
 Db 1497 cctctgcgtcgtcagccgaactataaaaggaactagtcctccggcaaaaaccccgtaa 1556
 |||
 QY 2898 ttgcagcagagagtagtggtggccggagcccgagagccagacccttctctccgg 2957
 |||
 Db 1557 ttgcagcagagagtagtggtggccggagcccgagagccagacccttctctccgg 1616
 |||
 QY 2958 gctcgcgagcagcagcggcgagcgtc 2984
 |||
 Db 1617 gctcgcgagcagcagcggcgagcgtc 1643
 |||

RESULT 4

AAH04729

ID AAH04729 standard; cDNA; 763 BP.

XX

AC AAH04729;

XX

26-JUN-2001 (first entry)

XX

Human cDNA clone (5'-primer) SEQ ID NO:1564.

XX

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

Homo sapiens.

XX

EP1074617-A2.

XX

07-FEB-2001.

XX

28-JUL-2000; 2000EP-0116126.

XX

29-JUL-1999; 99JP-0248036.

XX

27-AUG-1999; 99JP-0300253.

XX

11-JAN-2000; 2000JP-0118776.

XX

02-MAY-2000; 2000JP-0183767.

XX

09-JUN-2000; 2000JP-0241899.

XX

(HELI-) HELIX RES INST.

XX

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

PI

WPI; 2001-318749/34.

XX

Primer sets for synthesizing polynucleotides, particularly the 5602

XX

full-length cDNAs defined in the specification, and for the detection

XX

and/or diagnosis of the abnormality of the proteins encoded by the

XX

full-length cDNAs -

XX

Claim 1; SEQ ID 1564; 2537pp + CD ROM; English.

XX

The present invention describes primer sets for synthesizing 5602

XX

full-length cDNAs defined in the specification. Where a primer set

XX

comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX

to the complementary strand of a polynucleotide which comprises one of

XX

the 5602 nucleotide sequences defined in the specification, where the

XX

oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX

of an oligonucleotide comprising a sequence complementary to the

XX

complementary strand of a polynucleotide which comprises a 5'-end

XX

sequence and an oligonucleotide comprising a sequence complementary to a

XX

polynucleotide which comprises a 3'-end sequence, where the

XX

oligonucleotide comprises at least 15 nucleotides and the combination of

XX

the 5'-end sequence/3'-end sequence is selected from those defined in

XX

the specification. The primer sets can be used in antisense therapy and

XX

in gene therapy. The primers are useful for synthesizing polynucleotides,

XX

particularly full-length cDNAs. The primers are also useful for the

XX

detection and/or diagnosis of the abnormality of the proteins encoded by

XX

the full-length cDNAs. The primers allow obtaining of the full-length

XX

cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX

AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

XX

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

Query Match 10.4%; Score 336; DB 22; Length 763;
 Best Local Similarity 100.0%; Pred. No. 9.1e-62;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2896 aattgcagcagagtagtggtggccggagcccgagagccagacccttctctccc 2955
 |||
 Db 1 aattgcagcagagtagtggtggccggagcccgagagccagacccttctctccc 60
 |||
 QY 2956 gggctcgcgagcagcagcggcgagctccgcgcaacaacagagcggcttctcagggcgc 3015
 |||
 Db 61 gggctcgcgagcagcagcggcgagctccgcgcaacaacagagcggcttctcagggcgc 120
 |||
 QY 3016 ttgtctctgttttttcccggttctgttttctcccttctccggaggtctgtcaagg 3075
 |||
 Db 121 ttgtctctgttttttcccggttctgttttctcccttctccggaggtctgtcaagg 180
 |||
 QY 3076 ggtaggagaaagagacgcacacacaaagtggaacacagagtaagaggtctcagtgact 3135
 |||
 Db 181 ggtaggagaaagagacgcacacacaaagtggaacacagagtaagaggtctcagtgact 240
 |||
 QY 3136 tacttgggcgttattgtttgttcgagcccaagaggttcgggaaggtctcgggtttcg 3195
 |||
 Db 241 tacttgggcgttattgtttgttcgagcccaagaggttcgggaaggtctcgggtttcg 300
 |||
 QY 3196 gggactttgatccggagcccccacatccccaccactt 3231
 |||
 Db 301 gggactttgatccggagcccccacatccccaccactt 336
 |||

RESULT 5

AAH17451

ID AAH17451 standard; cDNA; 1750 BP.

XX

AC AAH17451;

XX

26-JUN-2001 (first entry)

XX

Human cDNA sequence SEQ ID NO:16905.

XX

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

Homo sapiens.

XX

EP1074617-A2.

XX

07-FEB-2001.

XX

28-JUL-2000; 2000EP-0116126.

XX

29-JUL-1999; 99JP-0248036.

XX

27-AUG-1999; 99JP-0300253.

XX

11-JAN-2000; 2000JP-0118776.

XX

02-MAY-2000; 2000JP-0183767.

XX

09-JUN-2000; 2000JP-0241899.

XX

(HELI-) HELIX RES INST.

XX

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

PI

WPI; 2001-318749/34.

XX

Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs -

XX

PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 9030; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AB11004-AB21534) and proteins
 CC (AB11478-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: the sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 5076 BP; 1435 A; 1052 C; 1191 G; 1398 T; 0 other;

Query Match 10.2%; Score 330.6; DB 22; Length 5076;
 Best Local Similarity 65.0%; Pred. NO. 2.1e-60;
 Matches 587; Conservative 0; Mismatches 299; Indels 17; Gaps 6;
 QY 375 ggccggcacaatggctgtaatcccgagcacttggaggcctgagccgagga 434

Db 3222 GGCCTGGTCATTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCTGAGCGCGGA 3163
 QY 435 tcgcttgagctccaagatttgagaccagccttgataacatggcaaaacccctgctctaca 494
 Db 3162 TCACCTGAGGTCAGGAGTTTGGAGCAAGCCTACCCAACTGGTGAACCC-CCATCTCTTA 3104
 QY 495 aaaaaatacaaaaattagatgggtgggtggcatgcaactgtggtcccagctacttggg 554
 Db 3103 CTAAATAACAAAATTAGCCAGGCATGGTGGCAGCGCTGTAAATCCAGCTACTCAGG 3044
 QY 555 aggttaaggtgggagatcgcttgagccagggagtgcaagtctacactgagccatattg 614
 Db 3043 AGGCTGAGGCAGAGAAATTCCTTGAACCCAGAGGTTGGAGGTCCAGTGCAGCGAGATCA 2984
 QY 615 gatcactgcaactccagcctg---ggtagacagagcagaccctgtctcaaaaaaagaaa 671
 Db 2983 CGGCATGTCATTCAGCTGGATAGATAGAGCAGAGACTCTGTCTCAAAAAAAGAGAG 2924
 QY 672 tgaagagaaaagaaagagagagagagatgaggggagagagagggggggggag 731
 Db 2923 AGAGAGAAAAAAGAGCTTTGATAAGGAACCTTGGCATAACCAATAATTTGGAAGATAA 2864
 QY 732 gaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 791
 Db 2863 AAGTCAAGGAAGCTTTCAGGATCTTTTACCAATTAATTTTAAAGTCTAAAGTGAACAACCTA 2804
 QY 792 ggcagaaagactttacgtataatgctcatcatgtggtgtgcaagtttgaccccaaaccc 851
 Db 2803 GCTAGCCAAATGAGAAATGCTGATTAATGGACTTATCAGGCTTAAGACT--GCCTTAAGCTG 2747
 QY 852 aattattgaccaggttattctttgactgaggaaggggggggggggggggggggggggggg 911
 Db 2746 CTGGGTAGAAATCAGGGGCACCTTGGCTTTCCCGTGGAGGGTCTTCAGGTGACGTTCTTATCA 2687
 QY 912 ggccttagaagctcatctctggcctttctgagatccatcccttcttttttttttttttt 971
 Db 2686 ACTTCTGACCAACCTGGCCCAATGCTTCCCTGGAAATAATTTTCTTTTCTTTTCTTTT 2627
 QY 972 gacacgagctctgctctgctcactcaggtggagtgagtgagtgagtgagtgagtgagtgag 1031
 Db 2626 GAGATGGAGTCTGCTCTGTCCACCCAGGCTGGAATGCAACGGTACCATCTCAGTCTACTG 2567
 QY 1032 taacctgtgctcccggttcaagcgattctctgctgactgactgactgactgactgactgact 1083
 Db 2566 CAACCTCTGCTCCCGGTTCAAGCGATCTCTGCTCCCTCAGCTCTCTGAGTAGTGGGAT 2507
 QY 1084 aacagcgcccgccac 1143
 Db 2506 TACAGGATGTGCCACCCAGCGGCTAAATTTTGTGA-TTTTAGTAGAGATGGGTTTCG 2448
 QY 1144 tcatgttggccaggttgggttcgaactcctgactgaggtgagctgcccacaccttgccctc 1203
 Db 2447 CCATGTTGGCCAGGCTGATCTCGAACTCTCGACCTCAGCTGATCTGCTGCTTGGCCTC 2388
 QY 1204 ccaag-tgctggattacaggaatgagcactgcccagcagcagcagcagcagcagcagcag 1262
 Db 2387 CCAAAGTGGAGGATACAGGCATGAGCCACCGCTGCCAGACAAATTTTTTTTTTTTTTTT 2328
 QY 1263 aag 1265
 Db 2327 AAG 2325
 RESULT 8
 ABA16701/c
 ID ABA16701 standard; DNA; 5075 BP.
 XX
 AC ABA16701;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 9032.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
OS
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220863.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228927.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0242221.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-024474.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.

PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231245.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232081.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-541565/60.
 XX
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 XX PS Disclosure; SEQ ID NO 11476; 1701pp + Sequence Listing; English.
 XX
 XX CC The invention relates to novel genes (AB11004-ABA21534) and proteins
 CC (AB14678-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SQ Sequence 23934 BP; 7785 A; 4761 C; 4365 G; 7023 T; 0 other;

Query Match 9.9%; Score 320.4; DB 22; Length 23934;
 Best Local Similarity 64.2%; Pred. No. 4.6e-58;
 Matches 592; Conservative 0; Mismatches 296; Indels 34; Gaps 6;
 QY 348 tctctccatttaagatgaagaacaggccggcacaatgctgctgtaattccca 407
 Db 21341 TTTCTCACATAAAGACCTTACTTAGGAGGCCAGAGCTGGTGCATGCTATAATCCCA 21282
 QY 408 gcactttggaggctgagccagagatcgcttgagctccagattagaccagcctgg 467
 Db 21281 GCACITTTGGAGGCCGAGGCTGGCAGATCGCTTGAGTGCAGGATTTGAGACGACATGG 21222
 QY 468 ataacatggcaaaacccctgtctctacaaaaatacaaaaaattagatgggtgtgtggc 527
 Db 21221 GCAACATGGCGAAACCCCATCTCTACAAAAATA---CAAAATTAACACAGGTTGTGGC 21165
 QY 528 atgcaacctgtgtcccaagctacttggaggcttaagtggaggatgcgttgaccacagg 587
 Db 21164 ACATACCTGTAGTCCAGCTACTGAGGAGGCTGAGGCAACAAGAAATGCTTGGACCTGGGA 21105

QY 588 agtcaagtctacagctagcagctgattgatcactcagctcagctggttagacagaca 647
 Db 21104 GCGGAGGTGTCAGTGAGCGAGATCATGCCACTGCATCTAGCCTGGG-CGACAGAGCA 21046
 QY 648 agacctgtctcaaaaaaagaaatgaagaagaaagaaagaaagagagagagagaga 707
 Db 21045 AGACTCTGTATCAAAAAAAAAAAAAAAAAAAAAAAAAAACAACACTAGGAACCTCTAAGTCAA 20986
 QY 708 tgagggggagggaggggggaggaaggaaggaaggaaggaaggaaggaaggaagga 767
 Db 20985 GTAGAAGTGATCAGATATTTATCTGCATAAATAATAAAGAGGTAGAAAGTAATTCATATTT 20926
 QY 768 aagaaaaaacaagatg-----aaacagaggcagaagactttacgaaattgctcat 820
 Db 20925 TAAGAAGGAACAATTTGGGTTTCATATCATGAAGACTTAATAGAGTTCAAAGAGGGA 20866
 QY 821 catgtggttgcaggttgaccaccaaacccaat-----ttattgaccaaggtta 870
 Db 20865 CATAAATTTCTACTTTGGGAGACAGAAAAATCATGGAATGTTCTTTGAAAAATTTT 20806
 QY 871 ttcttgactgaggaagggggtccgctctctctcctggttggctttgaaagactcatct 930
 Db 20805 ATTACAGATGAAGCCAGAAGAACTCTGGGTTGAGGAGAAGGCATTGGGAGCCAGCGG 20746
 QY 931 ctggccttctgagatccatcccttttttttttttttttttttttttttttttttttt 990
 Db 20745 TAGGGACTGGAGAGATCTG-----GAAGCATAAAGAAGAAATGAGACAGAGTATCACTCTG 20691
 QY 991 tcaactcagctgagtgcaagtgcagtgatctcgcactcactgaacctcgtcctcccggt 1050
 Db 20690 TCACCCAGGCTGGAGTGCGTAGACAACTCTGGCTACTGCAACCTCCACCTCCTCTGGGT 20631
 QY 1051 tcaagcattctcctgcctcagctc-----ctgagataaacaggccgcgcaccac 1102
 Db 20630 TCAAGCGATTCTCATGCTCAGCCTCCCAAGTAGCTGGATTACAGGCGCATGCCACGAT 20571
 QY 1103 atctggctaatttttatt 1162
 Db 20570 GCCTGGCTAAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTT 20511
 QY 1163 ttgcgaactcctgagtgagctgcccacttggcctcccaagtgtggaattaca 1222
 Db 20510 CTCCAATTCCTGACCTCATGTGATCCACCACCTCGGCTCCCAAGTGTGGGATTACA 20451
 QY 1223 ggcagtgagcactgcgcagc 1244
 Db 20450 GGCATGAGCCATCACACCGGC 20429
 RESULT 11
 AAL36171
 ID AAL36171 standard; DNA; 23934 BP.
 AC AAL36171;
 XX
 XX 08-JAN-2002 (first entry)
 XX
 XX Human musculoskeletal system related polynucleotide SEQ ID NO 2536.
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antinflammatory; antiulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200155367-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01338.

XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0234000.
 PR 14-SEP-2000; 2000US-0234240.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.

Db 3190 tagggactgagagatctg-----gaagcataaagaaggaatgagacagagatcactctg 3244
 QY 991 tcaactcagctgagtgagtgatctcgaactcactgaactcactgaactcctcccggt 1050
 Db 3245 tcaccagctgagtgagtgatctcgaactcactgaactcactcctcccggt 3304
 QY 1051 tcaagcattctcctcctcagctc-----ctgagataacagcgccgcccaccac 1102
 Db 3305 tcaagcattctcctcctcagctc-----ctgagataacagcgccgcccaccac 3364
 QY 1103 atctgctaatcttctgatttttagtaagactggttttcatctgttgccaggttgg 1162
 Db 3365 gctcgttaattttgatttttagtagagacaggttttaccatgttgccaggtgat 3424
 QY 1163 ttcgaactcctgacctgaggtgagctgcccacttggcctcccaagtctgggattaca 1222
 Db 3425 ctcaaatctcctgacctcatgtgatccaccactcggcctcccaagtctgggattaca 3484
 QY 1223 ggcagcactgcccagc 1244
 Db 3485 ggcagcactcaccacccg 3506

RESULT 12
 AAL36179/C
 ID AAL36179 standard; DNA; 23934 BP.
 XX
 AC AAL36179;
 XX
 DT 08-JAN-2002 (first entry)
 XX
 DE Human musculoskeletal system related polynucleotide SEQ ID NO 2544.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system; ds.
 OS Homo sapiens.
 XX
 PN WO20015367-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01338.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234984.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.

AA04522
ID AAL04522 standard; DNA; 23934 Bp.
XX
AC AAL04522;
DT 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 7210.
DE
XX Human reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
XX WO200155320-A2.
PN
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
PF
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184564.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234977.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.

PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225457.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-476224/51.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the respiratory system including respiratory cancers and also for testing and detection e.g. diagnosis -

Disclosure; SED ID No 777; 546pp; English.

The present invention relates to the isolation of novel human respiratory antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for these polypeptides. The sequences of the invention are useful for preventing, treating and/or prognosing disorders related to the respiratory system including throat disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of the respiratory tissues e.g. lung cancer. The polynucleotide sequences of the invention are useful in gene therapy and antisense therapy. AAS28161-AAS28764 represent genomic sequences encoding for novel human respiratory antigens.
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 23934 BP; 7785 A; 4761 C; 4365 G; 7023 T; 0 other;

Query Match 9.9%; Score 320.4; DB 22; Length 23934;

Best Local Similarity 64.28; Pred. No. 4.6e-58;

Matches 592; Conservative 0; Mismatches 296; Indels 34; Gaps 6;

Qy 348 tctctccattttaagatgaagaacagcgccgacaaatgctgaatgcctgtaattccca 407
 Db 21341 TTCTTCACATAAAGACCTTACTTAGGAGCCGAGACGCTGGTGCATATGCCTATATATCCCA 21282
 Qy 408 gcaactttggaggctgagccagaggtcgcttgagctccagagcttgagacagcctgg 467
 Db 21281 GCACCTTTGGAGCGCGAGCTGGCAGATCGCTTGAGTCGAGGAGCTTGAGACAGCATGG 21222
 Qy 468 ataatatgcaaaacccctgtctctacaaaaaaatacaaaaaatagatggtgtggtggc 527
 Db 21221 GCACATGGCGAAACCCCATCTCTACAAAAAATA---CAAAATTAACACAGGTGTGTGGC 21165
 Qy 528 atgacactgtgtcccaactgacttgagctgaagctgaagctgaagctgagcccaagg 587
 Db 21164 ACATACCTGTGTAGTCCAGCTACTGAGGAGCTGAGGACACAGAAATGCTTGAGCCTGGGA 21105
 Qy 588 agtcaagtctacactgagccatgattggtatcactgcactccagcctgggtagacagagca 647
 Db 21104 GCGGAGGTTGACGTGAGCCGAGATCATGCCACTGCATCTAGCCTGGG-CGACAGAGCA 21046
 Qy 648 agaccctgtctcaaaaaaaagaataaagaagaagaagaagaagaagaagaagaaga 707
 Db 21045 AGACTCTGTATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 20986
 Qy 708 tgaggag 767
 Db 20985 GTAGAAGTGATCAGATATTTATCTGCTATAATAATAAAGGTAGAAAGTAAATCATATTT 20926
 Qy 768 aagaaaaaaacaagatg-----aaacagagcagagaagaacttactgaattgctcat 820
 Db 20925 TAAGAGGAACAATTTGGGTTTCATACATCAATGAAGAGTTAAATGAGTTCAAAAGAGGGA 20866
 Qy 821 catgtggtgtgaagtttgaccccaaaacccaat-----ttattgaccaaggtta 870
 Db 20865 CATAAAATTTTCTACTTTGGGAGACAGAAATAATCATGGAATGTTCTCTTGAAAAATTTT 20806
 Qy 871 tctttgactgagcagaggggtccgtctctctggtggtggtggtggtggtggtggtggt 930
 Db 20805 ATTGAGAGATGAGCAGAGAGAACTCTGGTTGAGGAGAGGCAATGGGGAGCCAGCGG 20746
 Qy 931 ctggccttctgagatccatcccttcttcttcttcttcttcttcttcttcttcttctg 990
 Db 20745 TAGGACTGGAGAGATCTG-----GAAGCATAAAGAAAGATGAGACAGAGTATCACTCTG 20691
 Qy 991 tcaactgagctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1050
 Db 20690 TCACCCAGGCTGGAGTGCAGTAGCACAATCTTGCTCACTGCAACCTCCACCTCCTGGGT 20631
 Qy 1051 tcaagcgtattctctgctcagctc-----ctgagataacagcgcgcgcgcacccac 1102
 Db 20630 TCAAGGATTCATGCTCAGCTCCAGTACCTCCAGTACCTGGGATTACAGCGGATGCCACAT 20571
 Qy 1103 atctggttaatttttattttattttattttattttattttattttattttattttatt 1162
 Db 20570 GCCTGGCTAAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTT 20511
 Qy 1163 ttcaactctgagctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1222
 Db 20510 CTCCAATTCCTGAGCTCATGTGATTCACCCACCTCCGCGCTCCCAAGTGTGGGATTACA 20451
 Qy 1223 ggcagtgagcactgc 1244
 Db 20450 GGCATGAGCCATCACCACCGGC 20429

RESULT 15
 AAK71442
 ID AAK71442 standard; DNA; 23934 BP.
 XX
 AC AAK71442;

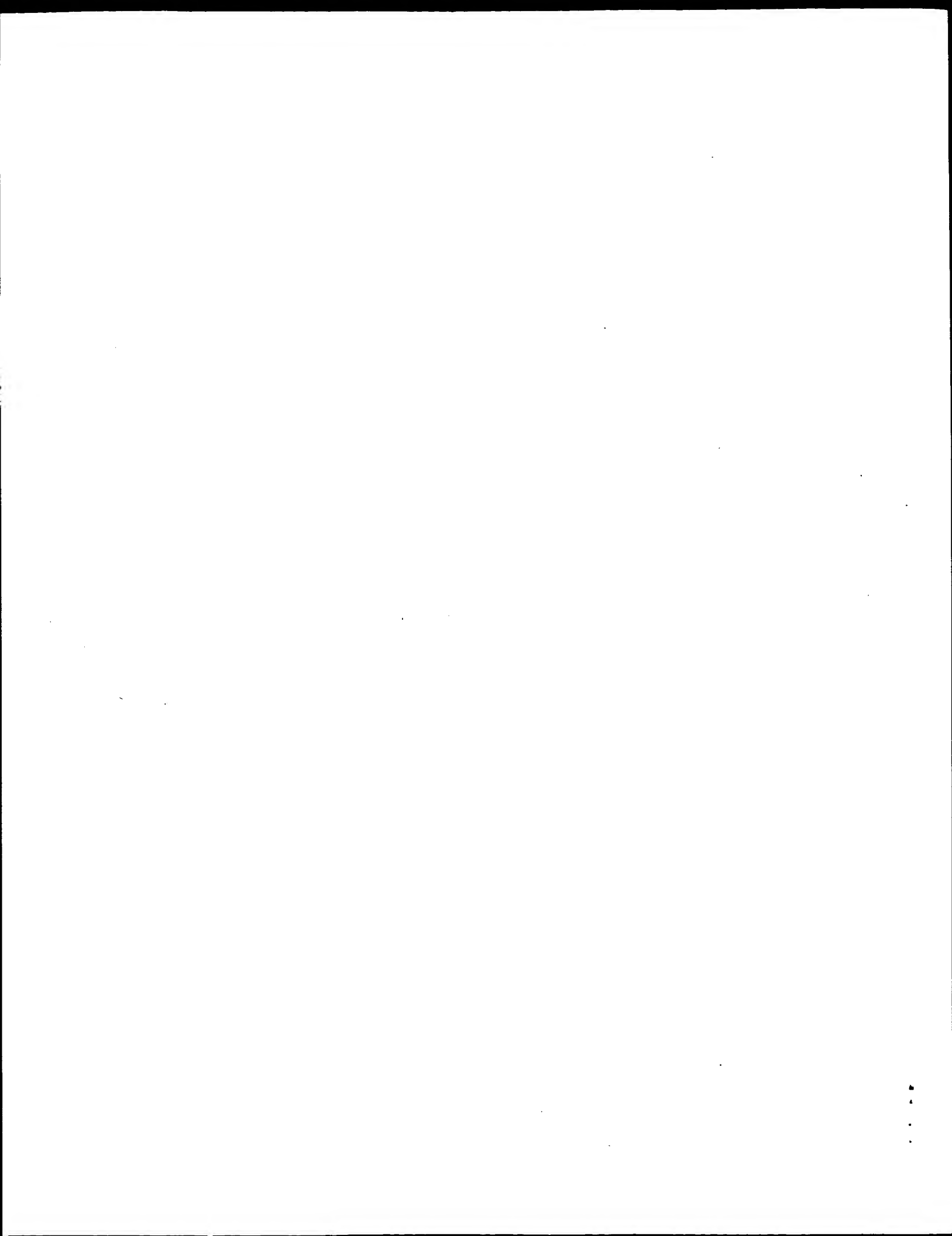
XX DT 06-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26254.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 KW Homo sapiens.
 OS
 XX WO200157182-A2.
 PN 09-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US01354.
 PF
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234984.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 26254; 3071pp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK51942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 23934 BP; 7023 A; 4365 C; 4761 G; 7785 T; 0 other;
 Query Match 9.9%; Score 320.4; DB 22; Length 23934;
 Best Local Similarity 64.2%; Pred. No. 4.6e-58;
 Matches 592; Conservative 0; Mismatches 296; Indels 34; Gaps 6;
 QY 348 tctctccatttttaagatgaagaaacaggccggccacaaatggcctaattgcctgaatccca 407
 Db 2594 ttctcacataaagaccctacttagaggccagagcgtgtgctcatcctataatccca 2653
 QY 408 gcactttggaggctgagccagaggatcgcttgagctccagagtttgagaccagctgg 467
 Db 2654 gcactttggaggccgaggtggtgcagatcgcttgagtcgaggagtttgagaccagatgg 2713
 QY 468 ataatatggcaaaacccctgtctctcaaaaaaaatacaaaaaatagatgggtgtgtgc 527
 Db 2714 gcaacatggcgaaccccatctctcaaaaaata---caaaatcaaccaggtgtgtgtgc 2770
 QY 528 atgcacctgtgtccagctacttggagggttaaggtggaggatcgcttgagccaggg 587
 Db 2771 acataacctgtagtcctactgagggaggtgagcacaagaattgcttgagcctgga 2830
 QY 588 agtcaagtctacatgagccatgatttgatcactgcactccagcctggtgacagagca 647
 Db 2831 ggcggaggttgagtgccgagatcatgcccactgcacttagctagggaacttaagtaa 2889
 QY 648 agacctgtctcaaaaaaaagaaatgaagaaagaaagaaagagagagagagagaga 707
 Db 2890 agactctgtatcaaaaaaaagaaagaaagaaagaaagaaagagagagagagagaga 2949
 QY 708 tgaggggaggggggggggaggaaggaaggaaggaaggaaggaaggaaggaagga 767
 Db 2950 gtagaagtgcagatatttcttcgcataataataaagtagaagtagaagtagaattt 3009
 QY 768 aagaaaaaacaagatg-----aaacagaggcgagaaagactttacgttaattgctcat 820

Db 3010 taagaaggaacaattggggtttcaactcatgaagagtttaaatgagttcaaaagagggga 3069
 QY 821 catgtggtttgaagtttgaccacaaacccaat-----ttattgaccacaggtta 870
 Db 3070 cataaaattttctacttttgaggagacagaaaaaatcatggaatgttctctgaaaaatttt 3129
 QY 871 ttctttgactgaggcaaggggtccgtctctctggtggttggccttttagaaaagctcatct 930
 Db 3130 atcagagatgaagccagaaagaaactctgggttgaggagaaggcattggggagccagcgg 3189
 QY 931 ctggcctttctgagatccatcccttcttcttttattttcttgacacgagctctgctctg 990
 Db 3190 tagggactggagagatctg----gaagcataagaagaaatgagacagagtatcactctg 3244
 QY 991 tcaactcaggttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1050
 Db 3245 tcaccaggttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 3304
 QY 1051 tcaagcattctcctgctcagcctc-----ctgagataacagcgccgccaccac 1102
 Db 3305 tcaagcattctcagctcagcctcagcctcagcctcagcctcagcctcagcctcagc 3364
 QY 1103 atctggctaattttgtatttttagtaagactgggttttcatcatgttgccaggttgg 1162
 Db 3365 gctggctaattttgtatttttagtagagacaggggttttaccatgttgccaggtgat 3424
 QY 1163 ttcgaaactcctgacctgaggtgagctgcccaccttgccctcccaaaagtctgggattaca 1222
 Db 3425 ctccaattcctgacctcagctgagctgcccaccttgccctcccaaaagtctgggattaca 3484
 QY 1223 ggcatgagccactgcgccagc 1244
 Db 3485 ggcatgagccatcacaccggc 3506

Search completed: September 20, 2002, 03:08:22
 Job time: 10571 sec



Page 4

Db	1318	AGCAGAGGACTGTGTCAGGCGCAGTCAAAATTTCTTCTGTATGATACTACAAATGTTAGACACATG	1255
QY	775	aaacaagatgaacacagaggcagaagactttacgtaaatgctcatcatggtggtgtcaaa	834
Db	1258	TCATTATACATTGTCCTCAAAATCTACAGAATGTACAAACCAAAAGCTTGGGTGATATGAG	1199
QY	835	gtttgacccccaaaacccaatttattgacccaaggttatctttgactgagggcaaggggtc	894
Db	1198	GAGTCAATGTTGGTCCATCAATTGCAACAAAAGGTACCCACAGTGGTGTAGGATGTGGATAA	1139
QY	895	cgtctctctggcccttgggctttagaagactcatctctggcccttctgagatccatccct	954
Db	1138	TGAGGAGGCTGTGCAGGTGTGGGCACAGTGGTATTAGAAATGC-----TCTATATTT	1084
QY	955	ttctttttattttttgacacggagtcttctctgtctactcaggtcgagtgacgtgc	1014
Db	1083	TCVTCTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	1024
QY	1015	atgatcgaactcactgaactctgctcccgggttcaagcaatctctcgcctcagcc	1071
Db	1023	ATGACTGTGGCTCACTGTAACTCCACTCCAGGTTCAAGAGATTTCTTCTGCTCAGCC	964
QY	1075	tcttgag-----ataacagggcgcgcacacacacatctggttaatttttattttta	1126
Db	963	TCTGAGTAGGTACAGTTTACAGGTGTGCACCACACCCAGCTAAATTTGTGTATTGTTA	904
QY	1127	gtaaaagactgggtttcatcatgttggccaggttgggttgcgaactctcgacctgaggtgag	1186
Db	903	GTAGACATGGGGTTTCATCATATTGGCCAGGCTGGCTTTAAACTCTCTGACCT--TCTGAT	846
QY	1187	ctgccacacttgccctcccaagtgcctgggtattacagcatgagccactgcgccagc	1244
Db	845	CTGCCACCTTGGCCTTCCAAAGTCTGGGATTACAGGTGTGAGCCACCCGCGCTGGC	788
RESULT 6			
US-08-724-394A-20/c			
; Sequence 20, Application US/08724394A			
; Patent No. 5872237			
; GENERAL INFORMATION:			
; APPLICANT: Feder, John N.			
; APPLICANT: Krommal, Gregory S.			
; APPLICANT: Lauer, Peter M.			
; APPLICANT: Ruddy, David A.			
; APPLICANT: Thomas, Winston			
; APPLICANT: Tsuchihashi, Zenta			
; APPLICANT: Wolff, Roger K.			
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1			
; TITLE OF INVENTION: Sequences and Antibodies thereto			
; NUMBER OF SEQUENCES: 31			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP			
; STREET: Two Embarcadero Center, 8th Floor			
; CITY: San Francisco			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94111-3834			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/724,394A			
; FILING DATE: 01-OCT-1996			
; CLASSIFICATION: 536			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Fitts, Renee A.			
; REGISTRATION NUMBER: 35,136			
; REFERENCE/DOCKET NUMBER: 017957-000100			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415-576-0200			
; TELEFAX: 415-576-0300			

[illegible]

```
QY 556 ggctaagtgaggagcgttgagcccgaggagtgcaagctctacacgtgagccatgattgg 615
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172415 GGCTGAGGCGAGAGATTGCTTGAACCTCGGAGGCGGAGTTGCAGCGAGCGGAGATTGC 172356

QY 616 atcaactgcacccagcctgggttagacagagcagaccctgtctcaaaaaaagaatgaa 675
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172355 GCCACTGCACCTCCAGCTGGTGACAGAGTGAGACTCCGCTCAAAAATAATAATAATA 172296

QY 676 agagaagaagaagaagagagagagagagagagagagagagagagagagagagagag 735
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172295 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 172236

QY 736 gaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 795
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172235 CTCTTGCGGTGTTTTCAGCGACTCTTCAGAGGCCCAAGAGTAAGTTTCCCTTAGCCC 172176

QY 796 gaaagactt----tacgtaaattgctcatctgtgttgctcaagtttgaccccaaaacc 851
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172175 CTACAGGTTCTTATGTTTAAATTTGTTACTCTCATTTAAGACATATTAATAATA 172116

QY 852 aattattgacaaagttattcttgaactgagggcagggggtccgctctcctgggacctg 911
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172115 CATGAGATTATTTCTGCATCATTTTGGTAAGATTGGCGGTTTCTCTTTGATCTC 172056

QY 912 ggtcttagaagctcatctctggcctttctgagatccatccctttcttttttttttt 971
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172055 TACTTCACACTGACCCACATATAAACATC- - -ACTGCCCTGTTTTTTTGTGTTGTTG 171999
```

RESULT 8

US-08-724-394A-22/c
Sequence 22, Application US/08724394A

Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Kronmal, Gregory S.

APPLICANT: Lauer, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

TITLE OF INVENTION: Sequences and Antibodies Thereto

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND and CREW LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/724,394A

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Fitts, Renee A.

REGISTRATION NUMBER: 35,136

REFERENCE/DOCKET NUMBER: 017957-000100

TELEPHONE: 415-576-0200

TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 246240 base pairs

TYPE: nucleic acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: misc_feature

LOCATION: 1..246240

OTHER INFORMATION: /note= "HLA-H.COMNTIG"

US-08-724-394A-22

Query Match 8.08; Score 257.8; DB 2; Length 246240;

Best Local Similarity 61.4%; Pred. No. 2.1e-53;

Matches 545; Conservative 0; Mismatches 312; Indels 30; Gaps 7;

QY 376 gccgggcacatggcctaagctgtaacccagcaccttggagggctgagccagagat 435

Db 172582 GCCGGGTGCAATGGCTACGCGCTGTAAATCCAGCAGCTTTGGGA -GCTGAGGCGGCGAC 172524

QY 436 gcgttgagctccagagtttgagaccagcctggataacacccctctctctacaa 495

Db 172523 CAC-----GAGGTCAAGACACACCCCTAGCGGAAATGGTAAACCCCTCTCTATT- 172473

QY 496 aaaaaatacaaaaattagatgggtgtgtggtgcatgcacctgtgtccagctacttggga 555

Db 172474 -AAAAATCAAAAGTGTACTGGCGGTGTGGCACATGCTGTAGTCCCGAGCTACTGGGA 172416

QY 556 ggctaagtgaggagatgccttgagccagggagtcgaagctacactgagccatgattgg 615

Db 172415 GGCTGAGGCGAGAGAAATGCTTGAACCTCGGAGGCGGAGGTTGCAGCGAGCGAGATTGC 172356

QY 616 atcactgcactccagcctgggttagacagagcaagaccctgtctcaaaaaaagaatgaa 675

Db 172355 GCCACTGCACCTCCAGCTGGTGACAGAGTGAGACTCCGCTCTAAAATAATAATAATA 172296

QY 676 agagaagaagaagaag 735

Db 172295 AT 172236

QY 736 gaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 795

Db 172235 CTCTTGCGGTGTTTTCAGCGACTCTTCAGAGGCCCAAGAGTAAGTTTCCCTTAGCCC 172176

QY 796 gaaagactt----tacgtaaattgctcatctgtgttgctcaagtttgaccccaaaacc 851

Db 172175 CTACAGGTTCTTATGTTTAAATTTGTTACTCTCATTTAAGACATATTAATAAGTG 172116

QY 852 aattattgacaaagttattcttgaactgagggcaaggggtccgctctcctgggacctg 911

Db 172115 CATGAGATTATTTCTGCATCATTTTGGTAAGATTGGCGGTTTCTCTTTGATCTC 172056

QY 912 ggtcttagaagctcatctctggcctttctgagatccatccctttcttttttttttt 971

Db 172055 TACTTCACACTGACCCACATATAAACATC- - -ACTGCCCTGTTTTTTTGTGTTGTTG 171999

APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match 7.8%; Score 253; DB 2; Length 246240;
Best Local Similarity 61.5%; Pred. No. 3.2e-52;
Matches 552; Conservative 0; Mismatches 315; Indels 30; Gaps 8;

QY 361 aagatgaagaaacagccgggcaaatgctgaatgctgaatccagcaatttgggaag 420
Db 171705 AAAAAAAAAACGGGGCGGCGCATGGCTCACTCTGTATCCAGCACATTTGGGAGG 171764

QY 421 ctgagggcagagatcgcttgagctccagagtttgagacagccttgataacatggcaaa 480
Db 171765 CTGAGGGGGGCCATCAC--GAGGTGAGGATCGAGACCATCTGTTAAACAAAGTGAA 171822

QY 481 accctgtctcaaaaaatacaaaaaatagatgggtgtgtgtgcatgcacctgtgtgt 540
Db 171823 ACCCGGTATCTAC--TAAAAATACAAAAATTAGCCGGGCTTGGTGGTGGCTGTAGT 171880

QY 541 ccagctacttggagcctaagtgagagatcgcttgagccaggaggaagctcaagctaca 600
Db 171881 CCAGCTGCTCAGAGGGCTGAGCGAGAGATGGCGGTGAATCCGGGAGCGGAGCTTGCA 171940

QY 601 ctgagccatgattggtcactgacgtccagcctgggtagacagagcaagaccctgtctca 660
Db 171941 GTGAGCGGAGATCACACCACCTGCACTCCAGCCTGGGCA-ACAGACCAAGACTCCGCTC 171999

QY 661 aaaaaaagaatgaagagagaaagaagaagagagagagagagagagagagagagagag 720
Db 172000 AAACAACAACAACAAAAACAGGCAGTGTATGTTTATGTTGGGTGCTAGTGTGAAGTAGA 172059

QY 721 gagggggggagggaagg---aagggaagggaaggaaaaaagaagataaagaaaaa 776

Db 172060 TCAAAGGAGAAACGCCCAATCTTTACCAATAATGATGTCAGAAATAATCTTCATGGAGA 172119
QY 777 acaagatgaacacagaggcagaaagactttacgtaaatgctcatcatgtgtgtccaagt 836
Db 172120 ACCCACTTAAATATGCTTAAATGAGAGTAACAATAATTAACATTAAGAACCTGTAGGGGC 172179
QY 837 ttgacccccaaaacccaatttattgaccaaaggttatctttgactgaggaagggggtccg 896
Db 172180 TAAGGGAAACCTTACTCTTTGGCCCTCTGAAGAGTGCCTGAAAAACCCACGACAGAGGAAG 172239
QY 897 ctctcctgggcttgggcttttagaaagctcatctctggcctttctgagatccatcccttt 956
Db 172240 ATTAATAGGATAAAATGATCCAATTTATTATTATTATTATTATTATTATTATTATT 172299
QY 957 ctt 1016
Db 172300 ATTATTATTATT-TTTAGCGGAGTCTCCTCTGTCAC-CAGGCTGGAGTGCAGTGGGCG 172357
QY 1017 gatctgactcactgtaacctctgctcccggttccagcgatctctcctgctcagcctc 1076
Db 172358 AATCTCGGCTGCTGCAACCTCCGCTCCCGAGTTCAAGCAATTTCTCTGCTCAGCCTC 172417
QY 1077 -----ctgagataacagcgcccgccgcacacatctggcttaatttttttttttttt 1128
Db 172418 CCCAGTAGCTGGGACTACAGGATGTGCCACCCAGCCAGCTAACTTTGTATTATTAA 172477
QY 1129 aaagactgggtttcatcatgttgccaggttggttgcgaactcctgacctgaggtgagct 1188
Db 172478 AGAGACGGGGTTTACCAATTTCCGGCTAGGGTGGTCTTGACCTCGTGGTCTG----- 172528
QY 1189 gccaccttggcctcccaagtgctgggattacaggeatgagccactgcgcccagct 1245
Db 172529 --CCGCGCTCAGCTCCCAAGTGTGGGATTACAGGCGTGAGCCATTGCACCCGCT 172583

RESULT 11
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200

[illegible]

RESULT 13

```

US-09-210-748A-3
: Sequence 3, Application US/09210748A
: Patent No. 6335156
: GENERAL INFORMATION:
: APPLICANT: Hermeking, Heiko
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kizler, Kenneth
: TITLE OF INVENTION: 14-3-3 SIGMA ARREST TH
: FILE REFERENCE: 1107.77810
: CURRENT APPLICATION NUMBER: US/09/210,748A
: CURRENT FILING DATE: 1998-12-15
: PRIOR APPLICATION NUMBER: 60/069,416
: PRIOR FILING DATE: 1997-12-18
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 7680
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-210-748A-3

```

Query Match

7.8%; Score 251.4; DB 4; Length 7680;

	Best Local Similarity	61.0%;	Pred. No. 1.5e-52;	
	Matches	575;	Conservative	0; Mismatches 286; Indels 82; Gaps
QY	371	aa	caggccgggcacaaatgcttaatccctgtaatcccgacactttgggagcgtgagccag	430
Db	5422	atctgccgggcattggtgcttatgctataatttcagcactttgggaagcctgagccag	5481	
QY	431	agatcgcttgagctccagagtttgagaccagcctggataacaatggcgaaccctgtctc	490	
Db	5482	tgtatcacttgagctcagctgtctcgagaccagcctggccaactggtgaaccacatctc	5541	
QY	491	tac-----aaaaaaatacaaaaattagatgggtggtggatgcacctgtgtgtcc	542	
Db	5542	tacttaaaaaaaatacaaaaattagctgggtgtggtgtgcacctgtgaatcc	5601	
QY	543	cagctacttgggaggtctaaggtygggagatcgtttgagcccgagggatcgaagtctacact	602	
Db	5602	cagctactcgggaggtcgaggcaggagaatcacttgaaccagaggtggaggttgcagt	5661	
QY	603	gagccatgattgatacactgcactcagctcagcctgggttagcacagacaagccctgtctcaaa	662	
Db	5662	gagctgagatgtgcacactcacactccagctgggtgacaagagcgaactccgtctcaaa	5721	
QY	663	aaaaagaaatgaagagaagaagaagaagagagagaggagagagatgaggggaggagga	722	
Db	5722	aaaaagaaaaataataataatgttg-----ccat	5754	
QY	723	ggggggggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga	782	
Db	5755	ccattctactgtctttccattaaactcgtgtaactccacaaagtcacactttatagttac	5814	
QY	783	tgaacacagagcgaagaagacttaagtaaaattgctcatcgtgtgtgtgctcaagtttgacc	842	
Db	5815	aggaactgaggtccacagagcttaatacacttggcca-----agg	5854	
QY	843	ccaaaacccaattattgaccaggttatctttgactgagcgaaggggtccgcctctcc	902	
Db	5855	ccaaaacagctataagaattacattttaggcgtctgattccaaagatactagtctattc	5914	
QY	903	tggccttggtttagaagctcatctctggtcccttctgagatccatcccttctctttt	962	
Db	5915	tgt-----atctcatagacaacaatacatattcaactttttgtgtgtg	5958	
QY	963	attttcttgacacgagctctgctgtcactcaggtcggagtcagtcggcatgatctc	1022	
Db	5959	tttgtttgagcggagcttgctctgtcccccaggtcggagtcagtcggcgcatctc	6018	
QY	1023	gactcaactgtaaactctgctccctccgggttcaagcgattctcgtcactcagctc-----	1076	
Db	6019	ggctcaactgcaacgtccgcctccgggttcaagcatctctcgtcagctcccgagtc	6078	
QY	1077	--ctgagataacaggccgcgcacacatcttggttaa--ttttgtatttttagtaaga	1133	
Db	6079	agctggactacaggcatgtgcccatgcccgctaatttttgtatttttagtagaga	6138	
QY	1134	ctgggttttcatcatgtttggcaggtttggtttcgaactcctgcactgaagtgagctgcccc	1193	
Db	6139	cagggtttctcgtggttagccagaatgggtctcgatctcctgcact--tgtgatccacca	6196	
QY	1194	cttggcctcccaaatgctgggattacaggcatgagccactggcccgccagctcagatcca	1253	
Db	6197	cctcagcttccccaaagtgctgagatgacagggcgtgagccaacgcgtccgacctattcca	6256	
QY	1254	tccctttctaaggccaacagctcca-tggtgcaaaaggggcca-tg	1296	
Db	6257	ctattttaaaattgagagataagaataaataa--tcaaaagggcgag	6299	

RESULT 14

US-08-257-963B-10 ; Sequence 10, Application US/08257963B ; Patent No. 5840686 ; GENERAL INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2002, 23:20:20 ; Search time 3895 Seconds
(without alignments)
11196.062 Million cell updates/sec

Title: US-09-846-456-1
Perfect score: 3231
Sequence: 1 acagggcatggggcagggtg.....gccccacatcccccaccactt 3231

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	10.4	763	9	AU121731
C 2	237.8	7.4	289	9	AW748338
C 3	221.8	6.9	736	9	AU135588
C 4	221.2	6.8	457	9	AW816516
C 5	219.4	6.8	577	12	AQ265389
C 6	218.6	6.8	544	12	AQ418551
C 7	215.2	6.7	719	12	AQ475181
C 8	214.2	6.6	581	12	AQ347610
C 9	213.4	6.6	338	9	AW023111
C 10	213.4	6.6	666	12	AG075934
C 11	213.2	6.6	493	12	B50449
C 12	213.2	6.6	622	12	AQ537948
C 13	213	6.6	514	9	AI754653
C 14	213	6.6	739	12	AQ200209
C 15	213	6.6	827	9	AV755654
C 16	212.8	6.6	484	12	AQ314669
C 17	212.8	6.6	487	12	AQ414845

18	212.4	6.6	501	12	AQ040260
19	212.2	6.6	416	9	AW341978
C 20	212.2	6.6	448	12	AQ634562
C 21	212.2	6.6	521	9	AW970571
C 22	212	6.6	575	12	AQ587429
C 23	211.4	6.5	577	9	AW973181
C 24	211.4	6.5	678	12	AQ387027
C 25	211.2	6.5	864	10	B6882030
C 26	211	6.5	1005	9	AL524675
C 27	210.8	6.5	458	9	AI733856
C 28	210.6	6.5	388	9	AW069227
C 29	210.6	6.5	437	9	AA644090
C 30	210.6	6.5	605	12	AQ412598
C 31	210.4	6.5	684	12	AG102895
C 32	210.2	6.5	924	10	B697612
C 33	210	6.5	477	12	AQ221138
C 34	210	6.5	577	9	AW719392
C 35	209.8	6.5	541	9	AW973187
C 36	209.8	6.5	671	12	AG050562
C 37	209.6	6.5	657	12	AG036967
C 38	209.4	6.5	461	12	B65075
C 39	209.4	6.5	642	12	B59854
C 40	209.4	6.5	832	10	B697217
C 41	209	6.5	483	9	AI275982
C 42	209	6.5	490	12	AQ426532
C 43	209	6.5	616	12	AQ554309
C 44	209	6.5	1750	11	AF119908
C 45	208.8	6.5	624	10	BF854308

ALIGNMENTS

RESULT 1

AU121731

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AU121731 763 bp mRNA linear EST 19-OCT-2000
AU121731 MAMMA1 Homo sapiens CDNA clone MAMMA1000851 5', mRNA
sequence.

ACCESSION AU121731

VERSION AU121731

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE Ota.T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,

JOURNAL Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

COMMENT Isoqai,T.

HRI human cDNA project

Unpublished (2000)

Contact: fakao isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3951

Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5' - & 3'-end one pass sequencing; Helix

Research Institute; CDNA library construction; Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1. 763

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MAMMA1000851"

/clone_lib="MAMMA1"

/tissue_type="mammary gland"

/note="Vector: pME18SFL3"

137 a 205 c 260 g 158 t

BASE COUNT 137 a 205 c 260 g 158 t

ORIGIN 3 others


```
/clone="2509010"  
/clone_lib="CITBI-E1"  
/sex="male"
```


Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES source
 Location/Qualifiers
 1. 719

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2589B9"
 /clone_lib="CITBI-El"
 /sex="male"
 /cell_type="sperm"

/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
 Caltech Human BAC Library D"
 BASE COUNT 228 a 162 c 180 g 149 t
 ORIGIN

Query Match 6.7%; Score 215.2; DB 12; Length 719;
 Best Local Similarity 75.7%; Pred. No. 5e-19;
 Matches 283; Conservative 0; Mismatches 83; Indels 8; Gaps 1;

Qy 893 tcgcgtctctggccttggttagaaagctcatctctggccttctctgagatccatcc 952
 Db 431 TTCCATCCTCAGAGCCCTGCGGCAACTGGTTTATCTGTGAATAATGGATTCCTTT 372
 Qy 953 cttcttt 1012
 Db 371 TTTTGTGTTTGTATTTTGTAGACAGATCTGTCTGTCACTCAAGCTGGAGTGCAGTG 312
 Qy 1013 gcatgatctcagctacgtatgaacctgtctcccggtttcaagcgattctctcctcag 1072
 Db 311 GCATGCTCTGCGCTCAGTCACTGCAACCTCCGCTCCGGGTTCAAGCGATTCTCTCCCTCAG 252
 Qy 1073 cctc-----ctgagataacagcgcccgccaccacatctgctgaatttttttttt 1124
 Db 251 CTCCTCCGGAACAGCTGGGATTACAGGCACCTGCCACCATGCTGCTATTTTGTATTTT 192
 Qy 1125 tagtaaacactgggttttctcatatgttgccaggttggttttcgaactcctgcactgaggtg 1184
 Db 191 TAGTAGAGACGGGGTTTGGCCATGTTGCCAGGCTGCTCGAACTCCTGACCTCAAGTG 132
 Qy 1185 agctgcccacactggcctcccaagtgctggattacaggcatgagcactgcgccagc 1244
 Db 131 APTCTCTCGCTCGGCTCCCAAGTGTGGGATTGCGAGGTGTGAGCCACCGCGCGCAGT 72
 Qy 1245 tcagatccatccct 1258
 Db 71 CTGGATCCATCTT 58

RESULT 8
 AQ347610/c
 LOCUS
 DEFINITION
 DNA sequence.
 ACCESSION
 AQ347610
 VERSION
 AQ347610.1
 KEYWORDS
 GSS.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 581)
 AUTHORS
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
 J.C.
 TITLE
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 JOURNAL
 Unpublished (1997)
 COMMENT
 Other_GSSs: RPCI11-126022.TV

Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES source
 Location/Qualifiers
 1. 581

/organism="Homo sapiens"
 /db_xref="GDB:7548357"
 /db_xref="taxon:9606"
 /clone="RPCI-11-126022"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"
 BASE COUNT 180 a 138 c 130 g 133 t
 ORIGIN

Query Match 6.6%; Score 214.2; DB 12; Length 581;
 Best Local Similarity 81.2%; Pred. No. 7.4e-19;
 Matches 264; Conservative 0; Mismatches 53; Indels 8; Gaps 1;

Qy 954 ttcttt 1013
 Db 454 TTCTTCTTATTTTTTTGTAGACAGATCTGTGTTGTACCCAGGTTGGAGTGCAGTG 395
 Qy 1014 catgatctcagctacgtatgaacctgtctcccggtttcaagcgattctctcctcagc 1073
 Db 394 CTGTAICTCGGCTCAGTCACTGCAAACTCTGCTCTGCTGCTCAAGCGATTCTCTCCCTCAGC 335
 Qy 1074 ctc-----ctgagataacagcgcccgccaccacatctgctgaatttttttttt 1125
 Db 334 CTCGAGTAGTGGGATTACAGCGCGCTGCCATCATGCTGGCTAATTTTGTATTTT 275
 Qy 1126 agtaaacactgggttttctcatatgttgccaggttggttttcgaactcctgcactgaggtga 1185
 Db 274 AGTAGAGATGGGGTTTTCCACCATCTGCGCAGATGGTCTCGAACTCCTGACCTCAGGTGA 215
 Qy 1186 gctgcccacacttgccctcccaagtgctgggattacaggcatgagcactgcgccagct 1245
 Db 214 TCTGCCACCTTGGCTCCCAAGTGTGGGATTACAGCGGTGAGCCACCGCCAGCC 155
 Qy 1246 cagatccatcccttttaaggga 1270
 Db 154 TTTATTGTTTATTCTTAGCTGAAA 130

RESULT 9
 AW023111/c

LOCUS
 DEFINITION
 IMAGE:2486822 5', mRNA sequence.

ACCESSION
 AW023111
 VERSION
 AW023111.1
 KEYWORDS
 EST.

SOURCE
 human.

ORGANISM
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 338)

AUTHORS
 Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.

Db	272	CTCAGGTGACCCCGCCCTCAGCCTCCCAAAATGCTGGATGTACAGGTGAGCCACCG	331
Qy	1237	cgccagctcagatccatcccttcttaaggga	1270
Db	332	CGCGCGCCCTCAGCCCTAACTTTCAATGACAA	365
RESULT	12		
LOCUS	AQ537948/c		
DEFINITION	622 bp DNA linear GSS 18-MAY-1999		
ACCESSION	AQ537948		
VERSION	AQ537948.1		
KEYWORDS	GI:4849638		
SOURCE	GSS.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 622)		
	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.		
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready		
JOURNAL	Map Building		
COMMENT	Unpublished (1997) Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbeetigr.org Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (inforesgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html . Seq primer: SP6 Class: BAC ends.		
FEATURES	Location/Qualifiers		
source	1..622		
	/organism="Homo sapiens"		
	/db_xref="GDB:7623762"		
	/db_xref="taxon:9606"		
	/clone="RPCI-11-323E19"		
	/clone_lib="RPCI-11"		
	/sex="Male"		
	/cell_type="Lymphocytes"		
	/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"		
BASE COUNT	123 a 178 c 103 g 218 t		
ORIGIN			
Query Match	6.6%	Score 213.2; DB 12; Length 622;	
Best Local Similarity	73.2%	Pred. No. 9.6e-19;	
Matches	300; Conservative	0; Mismatches 108; Indels 2; Gaps 2;	
Qy	361	aagatgaagaacagccggcgacaaatggctaagtcctgttaatccagcactttggagg	420
Db	462	AAATATTTCATGAGCGCGGCACAGTGGCTCACACCTGTAAATCGACACTTTGGGAGG	403
Qy	421	ctgaggccagaggatcgcttgagctccagagtttgagaccagccttgataacatggcaa	480
Db	402	CCAAAGGAGGAGAAATTTGCTTGGCCTCAGAGTTCAGGACCAAGCCCTGGGCAACATGGTGAG	343
Qy	481	acctgtctctacaaaaatacaaaaatagatgggtgtgtgtggcatgcacctgtggt	540
Db	342	ACCTCATCACATAAAATTAATTTAAAAAATAGCAGGACATGCTGGCATGCGCCTGCTG	283
Qy	541	cccagctacttggaggcgtaaggtggggaggatcgcttgagccagggagtcgaagtctaca	600

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:11:50 ; Search time 5225.75 Seconds
(without alignments)
1429.608 Million cell updates/sec

Title: US-09-846-456-2

Perfect score: 357

Sequence: 1 tggaggctcagctgagagg.....gagggaaggagctgtgttg 357

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_hgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB, ID	Description

1	357	100.0	357	6	AX351030	Sequence
2	353.8	99.1	149034	9	AF275948	Homo sapi
3	351.2	98.4	183999	6	AX092589	Sequence
4	340.2	95.3	480	9	HSA252277	Homo sapi
5	340.2	95.3	129608	9	AL353685	Human DNA
6	340.2	95.3	175064	2	AC012230	Homo sapi
7	340.2	95.3	201144	9	AF287262	Homo sapi
8	251.2	70.4	90698	2	AC021345	Homo sapi
9	184.2	51.6	200	9	AF258623	Homo sapi
10	159.4	44.6	10442	6	AX060713	Sequence
11	159.4	44.6	10442	6	AX060892	Sequence
12	159.4	44.6	10442	9	AF285167	Homo sapi
13	159.4	44.6	10474	6	AX060719	Sequence
14	159.4	44.6	10474	6	AX060721	Sequence
15	159.4	44.6	10474	6	AX060898	Sequence
16	159.4	44.6	10474	6	AX060900	Sequence
17	159	44.5	159	6	AX351033	Sequence
18	147	41.2	697	9	AF258627	Homo sapi
19	145.8	40.8	446	6	AX127764	Sequence
20	145.8	40.8	446	6	AX139751	Sequence
21	145.8	40.8	7260	6	AX253452	Sequence
22	145.8	40.8	9741	6	AX127830	Sequence
23	145.8	40.8	9741	6	AX139817	Sequence
24	145.8	40.8	9741	6	AX351038	Sequence
25	145.8	40.8	9854	6	AX127831	Sequence
26	145.8	40.8	9854	6	AX139818	Sequence
27	144.2	40.4	1556	9	AK024328	Homo sapi
28	136.8	38.3	298	9	AB037924	Homo sapi
29	135.8	38.0	7860	6	AX092594	Sequence
30	135.8	38.0	7862	6	AX135712	Sequence
31	101.8	28.5	207659	2	AC091466	Mus muscu
32	101.8	28.5	278572	10	AF287263	Mus muscu
33	99	27.7	99	6	AX351035	Sequence
34	99	27.7	99	6	AX351036	Sequence
35	67	18.8	7878	10	MMABC1	Mus musculu
36	65	18.2	6786	9	AB055982	Homo sapi
37	54.8	15.4	7074	5	AF362377	Gallus ga
38	42.6	11.9	22881	10	AF287142	Mus muscu
39	42.6	11.9	23024	6	AX080494	Sequence
40	42.6	11.9	185825	10	AC087114	Mus muscu
41	42.6	11.9	200143	2	AC073805	Mus muscu
42	39.4	11.0	226399	2	AC025794	Mus muscu
43	37.8	10.6	44762	2	AC011558	Homo sapi
44	37.6	10.5	55081	2	AC100427	Mus muscu
45	37.6	10.5	241192	2	AL603913	Mus muscu

ALIGNMENTS

RESULT	1	AX351030	357 bp	DNA	linear	PAT 06-FEB-2002
AX351030	Sequence 2 from Patent WO0183746.					
LOCUS	AX351030					
DEFINITION	Sequence 2 from Patent WO0183746.					
ACCESSION	AX351030					
VERSION	AX351030.1	GI:18616386				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (sites)					
AUTHORS	Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,					
	Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.					
TITLE	Regulatory nucleic acid sequences of the abcl gene					
JOURNAL	Patent: WO 0183746-A 2 08-NOV-2001;					
	Aventis Pharma S.A. (FR)					
FEATURES	Location/Qualifiers					
source	1. .357					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
BASE COUNT	64 a	78 c	116 g	99 t		
ORIGIN						


```

QY 1 tggaggtctcagctgagagggtgattagcagctcctcattgtgtatgggttcagca 60
Db 53130 TGGAGGTCTCAGCTGAGAGGGCTGGATTAGCACTCCTCATTTGGTGTATGGCTTTCAGCA 53189
QY 61 ataactgatggctgtttccctcctcctgctttatcttcttaagttaataaccagcagggcggt 120
Db 53190 ATAACGTGATGGCTGTTCCTCCCTCCCTTTCATTCAGTTAATGACCAACGACACACGCGT 53249
QY 121 ccctgctgcagctcctggtcgtctccaggggtcccgagccacacgctggggtgct 180
Db 53250 CCTGCTGTGAGCTCTGGCGGTGCTTCCAGGGCTCCGAGGCACACAGCTGGGGTGTCT 53309
QY 181 gctcagggaaacatggtcgttggcctcagctgaggttgcgtgtggaagaacotcact 240
Db 53310 GGCTGAGGAACATGGCTTGTGGCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACT 53369
QY 241 ttcaagaagaacacagtaagcttggtttttcagcagcgggggttctctcattttt 300
Db 53370 TTCAGNAGAGAACACAGTAAGCTTGGGTTTTTCAGCAGCGGGGGGTCTCTCATTTTTT 53429
QY 301 tcttgggtgtttgagttgggaggtggagggagggaggaagagcgtgtgtg 357
Db 53430 TCTTTGTGTTTTGAGTTGGGGATTGGAGGAGGAGGAGGAGGAAGAGCTGTGTG 53486

RESULT 4
HSA252277 HSA252277 480 bp DNA linear PRI 08-JAN-2001
LOCUS Homo sapiens partial ABC-1 gene for ATP-binding cassette
DEFINITION transporter-1, exon 2.
ACCESSION AJ252277.1 GI:12140344
VERSION ABC-1 gene; ATP-binding cassette transporter-1.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLES Porsch-Oezcuueruez M., Langmann, T. and Schmitz, G.
TRANSPORTER-1 (ABC-1) Promoter
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 480)
AUTHORS Porsch-Oezcuueruez, M.K.
TITLES Direct Submission
JOURNAL Submitted (07-JAN-2000) Porsch-Oezcuueruez M.K., Institute for
Clinical Chemistry, University of Regensburg,
Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY
FEATURES
source
1..480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/cell_type="leukocyte"
189..346
/gene="ABC-1"
/number=2
189..280
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
189..346
/gene="ABC-1"
281..>346
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
/codon_start=1
/product="ATP-binding cassette transporter-1"
/protein_id="CAC21428.1"
/db_xref="GI:12140345"
/translation="MACWPQLRLLLKLNLFRRRQT"
89 a 102 c 155 g 134 t
BASE COUNT
ORIGIN

```

```

Query Match 95.3%; Score 340.2; DB 9; Length 480;
Best Local Similarity 98.9%; Pred. No. 2.9e-93;
Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 tggaggtctcagctgagagggtgattagcagctcctcattgtgtatgggttcagca 60
Db 90 TGGAGGTCTCAGCTGAGAGGGCTGGATTAGCACTCCTCATTTGGTGTATGGCTTTCAGCA 149
QY 61 ataactgatggctgtttccctcctcctgctttatcttcttaagttaataaccagcagggcggt 120
Db 150 ATAACGTGATGGCTGTTCCTCCCTCCCTTTCATTCAGTTAATGACCAACGCGT 208
QY 121 ccctgctgcagctcctggtcgtctccaggggtcccgagccacacgctggggtgct 180
Db 209 CCTGCTGTGAGCTCTGGCGGTGCTTCCAGGGCTCCGAGGCACACAGCTGGGGTGTCT 268
QY 181 ggtcagggaaacatggtcgttggcctcagctgaggttgcgtgtggaagaacctcact 240
Db 269 GGCTGAGGAACATGGCTTGTGGCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACT 328
QY 241 ttcaagaagaacacagtaagcttggtttttcagcagcgggggttctctcattttt 300
Db 329 TTCAGNAGAGAACACAGTAAGCTTGGGTTTTTCAGCAGCGGGGGGTCTCTCATTTTTT 388
QY 301 tcttgggtgtttgagttgggaggtggagggagggaggaagcgtgtgtg 357
Db 389 TCTTTGTGTTTTGAGTTGGGGATTGGAGGAGGAGGAGGAGGAAGAGCTGTGTG 445

RESULT 5
AL353685/c AL353685 129608 bp DNA linear PRI 01-JUN-2001
LOCUS Human DNA sequence from clone RP11-31J20 on chromosome 9, complete
DEFINITION sequence.
ACCESSION AL353685
VERSION AL353685.23 GI:14329534
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLES Tracey, A.
JOURNAL Direct Submission
Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 8, 2001 this sequence version replaced gi:14272260.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-31J20 is from the library RPCI-11.1 constructed by the group

```

of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAGE3.6

IMPORTANT: This sequence is not the entire insert of clone
 RP11-31J20 it may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true right end of clone RP11-31J20 is at 129608 in this
 sequence. The true right end of clone RP11-413C10 is at 2000 in
 this sequence.

FEATURES

Location/Qualifiers

source

1. 129608
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-31J20"
 /clone_lib="RPC1-11.1"

repeat_region

1358..1653
 /note="AluSc repeat: matches 1. .288 of consensus"
 repeat_region
 1828..1877
 /note="L2 repeat: matches 2649. .2698 of consensus"
 repeat_region
 2496..2714
 /note="MIR repeat: matches 12. .250 of consensus"
 repeat_region
 2777..2896
 /note="MIR repeat: matches 6. .128 of consensus"
 repeat_region
 3237..3415
 /note="LIME repeat: matches 5696. .5821 of consensus"
 repeat_region
 6522..6818
 /note="AluSg repeat: matches 1. .295 of consensus"
 repeat_region
 7282..7415
 /note="LMB8 repeat: matches 6040. .6173 of consensus"
 repeat_region
 8145..8434
 /note="AluSc repeat: matches 1. .298 of consensus"
 repeat_region
 12145..12713
 /note="L2 repeat: matches 1363. .1940 of consensus"
 repeat_region
 13890..13969
 /note="L2 repeat: matches 2611. .2701 of consensus"
 repeat_region
 15380..15411
 /note="16 copies 2 mer ac 87% conserved"
 repeat_region
 16105..16144
 /note="10 copies 4 mer caca 100% conserved"
 repeat_region
 16868..17049
 /note="MIR repeat: matches 64. .246 of consensus"
 repeat_region
 17941..18229
 /note="AluSg repeat: matches 1. .287 of consensus"
 repeat_region
 18259..18553
 /note="AluSg repeat: matches 1. .293 of consensus"
 repeat_region
 20310..20616
 /note="AluSg repeat: matches 1. .308 of consensus"
 repeat_region
 20957..21107
 /note="MIR repeat: matches 49. .212 of consensus"
 repeat_region
 21783..22078
 /note="AluSx repeat: matches 7. .302 of consensus"
 repeat_region
 22320..22439
 /note="MIR repeat: matches 10. .146 of consensus"
 repeat_region
 22533..22839
 /note="AluSx repeat: matches 1. .307 of consensus"
 repeat_region
 23427..23945
 /note="L2 repeat: matches 2137. .2750 of consensus"
 repeat_region
 24245..24544
 /note="AluSg repeat: matches 2. .302 of consensus"
 repeat_region
 24556..24587
 /note="8 copies 4 mer acac 96% conserved"
 repeat_region
 26504..26561
 /note="29 copies 2 mer ta 69% conserved"
 repeat_region
 26849..26892
 /note="11 copies 4 mer tata 81% conserved"
 repeat_region
 27998..28103
 /note="MIR repeat: matches 28. .145 of consensus"
 repeat_region
 28515..28626
 /note="MIR repeat: matches 17. .129 of consensus"
 repeat_region
 28986..29213
 /note="MIR repeat: matches 2. .245 of consensus"
 repeat_region
 30121..30422

repeat_region
 31424..31734
 /note="AluSp repeat: matches 1. .302 of consensus"
 repeat_region
 31987..32116
 /note="MER5A repeat: matches 60. .188 of consensus"
 repeat_region
 32450..32536
 /note="MER5A repeat: matches 14. .106 of consensus"
 repeat_region
 33293..33380
 /note="22 copies 4 mer atgg 79% conserved"
 repeat_region
 34435..34588
 /note="MER5A repeat: matches 13. .189 of consensus"
 repeat_region
 34729..34873
 /note="MIR repeat: matches 49. .198 of consensus"
 repeat_region
 35802..35951
 /note="MIR repeat: matches 77. .248 of consensus"
 repeat_region
 37183..37260
 /note="2 copies 39 mer 92% conserved"
 repeat_region
 37673..37980
 /note="AluSc repeat: matches 1. .308 of consensus"
 repeat_region
 39674..40243
 /note="LIMD2 repeat: matches 5774. .6331 of consensus"
 repeat_region
 40256..40534
 /note="L2 repeat: matches 2256. .2533 of consensus"
 repeat_region
 41476..41615
 /note="MIR repeat: matches 30. .185 of consensus"
 repeat_region
 42010..42194
 /note="MIR repeat: matches 49. .234 of consensus"
 repeat_region
 44809..45239
 /note="LTR16A repeat: matches 6. .450 of consensus"
 repeat_region
 45359..45486
 /note="MIR repeat: matches 1. .139 of consensus"
 repeat_region
 46654..46693
 /note="10 copies 4 mer acac 97% conserved"
 repeat_region
 47118..47429
 /note="AluSx repeat: matches 1. .312 of consensus"
 repeat_region
 47915..48083
 /note="AluSg repeat: matches 145. .313 of consensus"
 repeat_region
 49097..49181
 /note="L2 repeat: matches 2625. .2710 of consensus"
 repeat_region
 49578..49758
 /note="MIR repeat: matches 9. .188 of consensus"
 repeat_region
 49847..49873
 /note="MIR repeat: matches 155. .182 of consensus"
 repeat_region
 49874..50188
 /note="AluJb repeat: matches 3. .311 of consensus"
 repeat_region
 50189..50347
 /note="MIR repeat: matches 5. .155 of consensus"
 repeat_region
 51677..51977
 /note="AluSx repeat: matches 1. .303 of consensus"
 repeat_region
 52025..52123
 /note="L2 repeat: matches 1655. .1753 of consensus"
 repeat_region
 54337..54544
 /note="MIR repeat: matches 25. .246 of consensus"
 repeat_region
 54754..55032
 /note="AluSp repeat: matches 1. .302 of consensus"
 repeat_region
 55042..55343
 /note="AluY repeat: matches 5. .298 of consensus"
 repeat_region
 56027..56232
 /note="MIR repeat: matches 21. .259 of consensus"
 repeat_region
 56233..56286
 /note="18 copies 3 mer tgt 72% conserved"
 repeat_region
 56342..56673
 /note="L1MA9 repeat: matches 5893. .6307 of consensus"
 repeat_region
 56750..56889
 /note="L2 repeat: matches 2565. .2710 of consensus"
 repeat_region
 57406..57567
 /note="MER5A repeat: matches 4. .187 of consensus"
 repeat_region
 57864..58174
 /note="AluSg repeat: matches 1. .311 of consensus"
 repeat_region
 59443..60004
 /note="L2 repeat: matches 2088. .2723 of consensus"
 repeat_region
 60201..60355
 /note="MIR repeat: matches 20. .183 of consensus"

```
repeat_region 63991..64159
/note="MIR repeat: matches 78..243 of consensus"
repeat_region 64590..64653
/note="32 copies 2 mer ta 70% conserved"
repeat_region 64593..64656
/note="16 copies 4 mer atat 68% conserved"
repeat_region 64981..65042
/note="MERSA repeat: matches 15..72 of consensus"
repeat_region 65043..65341
/note="AluX repeat: matches 1..299 of consensus"
repeat_region 65342..65441
/note="MERSA repeat: matches 72..186 of consensus"
repeat_region 66825..67131
/note="AluX repeat: matches 1..306 of consensus"
```

```
Query Match 95.3%; Score 340.2; DB 9; Length 129608;
Best Local Similarity 98.9%; Pred. No. 7.2e-93;
Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 tggagctcagctgagagcgtgattgagcagctcattgtgtgtgcttgcagca 60
|||||
Db 120580 TGGAGCTCAGCTGAGAGCGGTGGATTGACAGTCTCTATTGGTGTATGGCTTTCAGCA 120521
|||||

QY 61 ataactgagctgtttcccccctcctgctttattcttcagtttaataaccagccacggcgct 120
|||||
Db 120520 ATAACTGATGGCTGTTTCCCTCCTGCTTTATCTTCACTTAATGACCAAGCCAC-GGCGT 120462
|||||

QY 121 cctgtgtcagctgtgagcgtgcttcacagggctccagagccacagcgtggcgct 180
|||||
Db 120461 CCCTGTGTGAGCTGTGGCGCTGCTCCAGGGCTCCGAGCCACAGCTGGGGTGTCT 120402
|||||

QY 181 ggcctgaggaacatgcctgtgtgctcagctgagctgtgctgtggaagacacctcact 240
|||||
Db 120401 GGCCTGAGGAACATGCTGTGTGGCTCAGCTGAGCTGCTGCTGTGGGAACCTCACT 120342
|||||

QY 241 ttcaagaagaacacagtaagcttgggttttttccagcgcgggggttctctcattttt 300
|||||
Db 120341 TTCAAGAAGAACACAGTAAGCTTGGGTGTTTTCAGACGCGGGGGTCTCTCATTTT 120282
|||||

QY 301 tctttgtgtttgattggagattgagagggagggaggaagaaagaaagctgtgtg 357
|||||
Db 120281 TCFTTGTGTTTTGAGTTGGGGATTGGAGGGAGGGAGGAGGAAGAAAGCTGTGTG 120225
|||||
```

```
RESULT 6
AC012230/C 175064 bp DNA linear HTG 22-APR-2000
LOCUS Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
DEFINITION pieces.
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Collangelo,M., Collins,S., Collamore,A.,
Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
Galan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
```

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 22, 2000 this sequence version replaced gi:6454033.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2510

Center clone name: L.M.10

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 117571 bases at least Q40

Consensus quality: 145749 bases at least Q30

Consensus quality: 160940 bases at least Q20

Insert size: 185000; agarose-fp

Insert size: 171264; sum-of-contigs

Quality coverage: 2.9 in Q20 bases; agarose-fp

Quality coverage: 3.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 39 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1003: contig of 1003 bp in length

1004 1103: gap of 100 bp

1104 2634: contig of 1531 bp in length

2635 2734: gap of 100 bp

2735 4415: contig of 1681 bp in length

4416 4515: gap of 100 bp

4516 5785: contig of 1270 bp in length

5786 5885: gap of 100 bp

5886 7879: contig of 1994 bp in length

7880 7979: gap of 100 bp

7980 9686: contig of 1707 bp in length

9687 9786: gap of 100 bp

9787 12253: contig of 2467 bp in length

12254 12353: gap of 100 bp

12354 15228: contig of 2875 bp in length

15229 15328: gap of 100 bp

15329 17200: contig of 1872 bp in length

17201 17300: gap of 100 bp

17301 20131: contig of 2831 bp in length

20132 20231: gap of 100 bp

20232 22587: contig of 2356 bp in length

22588 22687: gap of 100 bp

22688 25707: contig of 3020 bp in length

25708 25807: gap of 100 bp

25808 26184: contig of 2377 bp in length

26185 28284: gap of 100 bp

28285 31338: contig of 3054 bp in length

31339 31438: gap of 100 bp

31439 34299: contig of 2861 bp in length

34300 34399: gap of 100 bp

34400 38318: contig of 3919 bp in length

38319 38418: gap of 100 bp

38419 42835: contig of 4417 bp in length

42836 42935: gap of 100 bp

42936 45448: contig of 2513 bp in length

45449 45548: gap of 100 bp

exon /number=18
140179. .140350
/gene="ABCA1"
exon /number=19
141340. .141471
/gene="ABCA1"
exon /number=20
142777. .142919
/gene="ABCA1"
exon /number=21
143123. .143260
/gene="ABCA1"
exon /number=22
143963. .144183
/gene="ABCA1"
exon /number=23
145441. .145513
/gene="ABCA1"
exon /number=24
146657. .146859
/gene="ABCA1"
exon /number=25
148527. .148575
/gene="ABCA1"
exon /number=26
148771. .148884
/gene="ABCA1"
exon /number=27
150280. .150428
/gene="ABCA1"
exon /number=28
152078. .152202
/gene="ABCA1"
exon /number=29
153438. .153536
/gene="ABCA1"
exon /number=30
156568. .156757
/gene="ABCA1"
/number=31

Query Match 95.3%; Score 340.2; DB 9; Length 201144;
Best Local Similarity 98.9%; Pred. NO. 7.7e-93;
Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 tggaggtctcagctgagagggctgattagcagctcattggtgtatggtgtttgcagca 60
|||||
Db 58220 TGGAGGCTCAGCTGAGAGGGCTGGATTAGCAGTCTCATTTGGTGTATGGCTTCGACGA 58279
|||||
QY 61 ataactgatgctgtttccctcctcgtcttatcttcagtttaataaccagccagcgcgt 120
|||||
Db 58280 ATAACTGATGGCTGTTTCCCTTCCTGCTTTATCTTTCAGTTPAATGACGACGAC 58338
|||||
QY 121 cctgtgctcagctgtgctgctcctccagggctccagccacagcgtggcggtgct 180
|||||
Db 58339 CCCTGCTGTGAGCTGTGCGCTGCTTCCAGGGCTCCGAGCCACACGCTGGGGTGTCT 58398
|||||
QY 181 ggctgagggaacatgagctgttgctcagctgaggtgtcgtcgtgtgaagaacctcact 240
|||||
Db 58399 GGCTGAGGGAACATGGCTTGTGGCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACT 58458
|||||
QY 241 ttacagagagacaacagctggttttttcagcagcgggggttctctcattttt 300
|||||
Db 58459 TTCAGAGAGACAAACAGTACGTGGGTTTTTCAGACCGGGGGTTCCTCAITTTT 58518
|||||
QY 301 tctttgtgtttttgagttggagtgagggagggagggaggaagagctgtgttg 357
|||||
Db 58519 TCITTTGTGTTTTTGTAGTTGGGATTGGAGGAGGGAGGGAGGGAAGGAGCTGTGTG 58575
|||||

RESULT 8
AC021345/c
LOCUS AC021345 90698 bp DNA linear HTG 13-JUL-2000

DEFINITION Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG: HTGS_PHASE0.
SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 90698)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens, clone RP11-24J9

Unpublished

2 (bases 1 to 90698)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,

Boguslavskiy,L., Boukhgalter,B., Brown,A., Buckett,G., Castle,A.,

Choepel,X., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,

Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenesstor,J.,

Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,

Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,

Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,

McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,

Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,

Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,

Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,

Zimmer,A. and Zody,M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6705761.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4483

Center clone name: 24_J_9

* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1873: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length
* 2825 2924: gap of 100 bp
* 2925 3802: contig of 878 bp in length
* 3803 3902: gap of 100 bp
* 3903 4816: contig of 914 bp in length
* 4817 4916: gap of 100 bp
* 4917 5759: contig of 843 bp in length
* 5760 5859: gap of 100 bp
* 5860 6764: contig of 905 bp in length
* 6765 6864: gap of 100 bp
* 6865 7747: contig of 883 bp in length
* 7748 7847: gap of 100 bp

Qy	241	ttcagaagaacacaaacagtaagcttgggtttttc-agcagcgggggtttctc	292
Db	12931	TTCAAGAAGACAAAACAGTAAGCTTGGGTTTTTCAAGCAGCAGCGGGGTTNTC	12879
RESULT 9			
AF258623S2			
LOCUS			
DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene, exon 2.			
ACCESSION AF258624			
VERSION AF258624.1 GI:7769714			
KEYWORDS			
SEGMENT			
SOURCE			
ORGANISM			
human.			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Auquierat,B.E., Fielding,C.J. and Kane,J.P.			
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence, promoter region, and four polymorphisms			
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press			
REFERENCE			
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Auquierat,B.E., Fielding,C.J. and Kane,J.P.			
TITLE Direct Submission			
JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA			
REMARK			
FEATURES			
source			
1..200			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosome="9"			
/map="9q31"			
22..179			
/gene="ABCA1"			
/number=2			
BASE COUNT 37 a 56 c 57 g 50 t			
ORIGIN			
Query Match 51.6%; Score 184.2; DB 9; Length 200;			
Best Local Similarity 98.0%; Pred. No. 1.6e-45;			
Matches 197; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy 79 cctctcgttattatcttctagtttaataacacagccacggcgctcctgctgctcagctctg 138			
Db 1 CCCTCCGCTGCTTAFCTTTCAGTTAATGACGACGAC-GGGCTCCCTGCTGAGCTCTGG 59			
Qy 139 ccgctgccttcacgggctcccgagccacagctggcgctgctggctgaggaacatggca 198			
Db 60 CCGTGCCTTCCAGGGCTCCGAGCCACACGCTGGGGGTGCTGGCTGAGGAACATGGCT 119			
Qy 199 tgttggcctcagctgaggttgctgctgtggaagaacaccttccactttcagaagaagacaaca 258			
Db 120 TGTTGGCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAACA 179			
Qy 259 gtaagcttgggtttttcagca 279			
Db 180 GTAAAGCTTGGGTTTTTCAGCA 200			
RESULT 10			
AX060713			
LOCUS			
DEFINITION AX060713 10442 bp DNA linear PAT 22-JAN-2001			
SEQUENCE 1 from Patent WO0078972.			
ACCESSION AX060713			
VERSION AX060713.1 GI:12406103			

Db 255 CCGGACCCACGCTGGGGCTGCTGGCTGAGGGAACATGGCTTGTGGCTCAGCTGAGG 314

Qy 217 ttgctgtgtggaagaacactcacttctcagaagaacacagtagtgggtttttca 276
 |||||
 Db 315 TTGCTGCTGTGGAAGAACCTCACTTTCAGAGAAGACAAATGTCAGCTGTACTGGAA 374
 |||||

Qy 277 gaagcgggggtctctctatcttcttctgtgtgtgttttgagttggg 321
 |||||
 Db 375 GTGGCCTGGCCTCACTTATCTCTCTGATCCCTGATCTGCTGTTCGG 419

RESULT 12

AF285167 AF285167 10442 bp mRNA linear PRI 09-AUG-2000

LOCUS Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA,
 complete cds.

ACCESSION AF285167

VERSION AF285167.1 GI:9755158

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 10442)

AUTHORS Schwartz,K., Lawn,R.M. and Wade,D.P.

TITLE ABCA1 gene expression and apoA-I-mediated cholesterol efflux are regulated by LXR

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 10442)

AUTHORS Lawn,R.M., Wade,D.P., Garvin,M.R., Wang,X., Schwartz,K., Porter,J.G., Seilhamer,J.J., Vaughan,A.M. and Oram,J.F.

TITLE Direct Submission

JOURNAL Submitted (06-JUL-2000) Discovery Research, CV Therapeutics Inc., 3172 Porter Drive, Palo Alto, CA 94304, USA

FEATURES

Location/Qualifiers

1..10442

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/map="9q31"

/cell_type="fibroblast"

/tissue_type="skin"

1..10442

/gene="ABCA1"

291..7076

/gene="ABCA1"

/codon_start=1

/product="ATP-binding cassette transporter 1"

/protein_id="BAF98175.1"

/db_xref="GI:9755159"

/translation="MACPQLRLLLKWLKNTLFRRTQCLLEAVMLPFIILISVRL
 SYPPYEQEHCFPNKAMPAGTLPLWVGIIICNANNPCFRYPPTGEAFGVGNFNKSIY
 ARLFSDARLLLYSOKDTSMKMRKVLRTLOQIKKSNLKLQDFLVNDNETFSGLYH
 NLSLPKSTVDKMLRADVTLHKVTLGGYOLHLSLCSNGSKSEMIQLGQVSELCGLP
 KEKLAARERLSNMDILKPLILVINSTPSPKELAEATYLLSHLSGLTAEQLFSMR
 SWSMDRQEWELTNVNSSSSTQIYQAVSRIVCGHGGKIKSLKWNIEDNNTKALF
 RGWTEAEATFYDNSTPYCNLDMLNLESSLRSRIWKALKPLLVKILYTPDTPTAT
 ROVMAAANKQIEALAFVHDLGEMMEELSPKIWTFMENSGMDLVRMLIDRNDHFE
 QOLGDLMTAODIVAFIAKHPEDVQSSNGSVYTWREAFNETNOAIRTSIREMECVNL
 KLEPTATVEVLINKSMELLDERKFWAGIVFTGTPGSTELPHVVKYKIRMDIDNVERT
 NKIKDGYNDPGRADPFEDMRKVGGRAYLDVVEQALIRVLTGTEKTKGVYMQMPY
 PCYVDIIFLRVNSRSMPLMTLAWIISVAVIIGIVIEKEARKETMRIGLNDLSILW
 FSWFTSSILPLVNSAGLLVILKLGNLPLPYDGKSVFVFSVFAVVTILOCFLLSTL
 SRANAACGIIYITFLPYLVCAWQDYVGTIKIPASLSPVAFGCEYFALFE
 EQGIGQWMDNLFESPEDEGFLNLTSSIMLFDFTFLXVMTWYIEAVFPQYGIPLPW
 YFPCFKSWFGESEDESKHSGSNOKRMSIECMEREPHLKLGVSIONLKVYRDGMKV
 AVDGLAFYEQGLTSLFGLHNGAKTMTSLTGLFPPTSGTAVILCKDILFSESTIR
 QNLGVCPOHNVFLDMLTVEEHWIYFARKLGLSEKRVRAEMEQMALDVLGSLSKLST
 SOLSGMDQRKILSVAFIISHGKLCVCGSSFLKQLGTGYTLTVLRKDVESLSGCRNS
 HMMDADYLDRIAIISHGKLCVCGSSFLKQLGTGYTLTVLRKDVESLSGCRNS
 STVYLKVEDSSVSSDAGLSDHESDTLTIDVSAISNLTIRKHFVSEARVLEIDGHEL
 TVPLVEAAKKGAFVFLPHEIDRLSDGISYISSETTLEIFLKVAEESGYDAETS
 DGLTPARNRRRAFGDKQSLRPFTEDDAADPNDSIDIPESRETDLLSGMDGKGYQVK

GWKLTQQQFVALLWKRLLIARRSRKFFAQIIVLPAVFVCIALVESLIVPPGKYPSLE
 LQPMYNEQTFVSDAPEDTGTLELLNALTIDKDPGFGTRCMENPIIDTPPCOAGEBW
 TTAPVPTIMDLFQNGNWTMNPSPACQSSDKIKKMLPVCPPGAGGLPPQKQNTA
 DILQDLTGRNISDYLVKTYVQIIIAKSKNKIWNNEFRYGFSLGVSNTQALPQREVN
 DAIKOMKHLKLAKDSSADRFNLNLSGRFMTGLDTRNNKVFNNKGMHAISSGFLNIN
 NATLRANLOGENPSHYGTTAFNHLNLTQOOLSEVALMTTSDVDLVSICVLFAMSEFV
 PASFVFLIQERSKAKHLQFISGVKPVIIWLSNFVDMCNVVPVATILVILFICFOQ
 KSYVSNLPVLALLLLYGSITPLMIPASFVKIPSTAYIVLTSVNLFIGINGSVA
 TFVLELTDNKLNNINDILKSVFLIPEHFCGLGLDMDYKNAMADALERFERREVS
 PLSWDLVRNLFAMAVEGVFFLITLQYREFIRPRPVNAKLSLUNDEDEVDREERQ
 RIDGGQNDILEIKELTKIYRKRKRPVDRICVGPPEGCEGLLVNGVAGKSSIFKFM
 LTGDTVTBGDAFLNKSILSNIRKHEPVQNMGYCPOFDAITELLTGREHVEFFALLRGV
 PEKEVKYGEWAIRKGLVKYGEKYGAGNYSNGNKRKLSAMALIGGPPVFLDEPTTG
 MDPKARFLWNCALSVVKGSRVSVLTSHSMECEALCTRMALMNGRFRCLGSVQHLK
 NREFGDGYTIVYRIAGSNPDLKPQDFGLAFPGSLVLEKHNMLQYLPSSLSLARI
 FSTLSQKKRLHIEDYSVSQTTLDQVFNFAKQSDDDHKLKDLSLHKNQTVVDVAVL
 SFLQDEKVKRESV"

BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others

ORIGIN

Query Match 44.6%; Score 159.4; DB 9; Length 10442;
 Best Local Similarity 81.8%; Pred. No. 1.2e-37;
 Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 97 cagtaataagaccagccagcggtccctgtgtgtcagctgtggtccgtgcttccagggct 156
 |||||
 Db 195 CAGTTAATGACACGACGGCGCTCCCTGTGTGAGCTCTGGCGCGTGCCTTCCAGGGCT 254
 |||||

Qy 157 cccagagccacgctggcggtgctgagggagacatggcatgttggtcctcagctgag 216
 |||||
 Db 255 CCCGAGCCACACGCTGGCGGTGCTGGCTGAGGGAACATGCTGTGCTCAGCTGAGG 314
 |||||

Qy 217 ttgctgtgtggaagaacactcacttctcagaagaacacagtagtgggtttttca 276
 |||||
 Db 315 TTGCTGCTGTGGAAGAACCTCACTTTCAGAGAAGACAAATGCTGACTTACTGGAA 374
 |||||

Qy 277 gaagcgggggtctctctatcttcttctgtgtgtgttttgagttggg 321
 |||||
 Db 375 GTGGCCTGGCCTCACTTATCTCTCTGATCCCTGATCTGCTGTTCGG 419

RESULT 13

AX060719

LOCUS Homo sapiens

DEFINITION Sequence 7 from Patent WO0078972.

ACCESSION AX060719

VERSION AX060719.1 GI:12406108

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 10474)

AUTHORS Lawn,R.M., Wade,D. and Garvin,M.

TITLE Regulation with binding cassette transporter protein abcl

JOURNAL Patent: WO 0078972-A 7 28-DEC-2000;
 CV THERAPEUTICS, INC. (US)

FEATURES

Location/Qualifiers

1..10474

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others

ORIGIN

Query Match 44.6%; Score 159.4; DB 6; Length 10474;
 Best Local Similarity 81.8%; Pred. No. 1.2e-37;
 Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 97 cagtaataagaccagccagcggtccctgtgtgtcagctgtggtccgtgcttccagggct 156
 |||||
 Db 227 CAGTTAATGACACGACGGCGCTCCCTGTGTGAGCTCTGGCGCGTGCCTTCCAGGGCT 286

QY 157 cccagaccacacgctggcgtgctgagggagaaacatggtgcatgttggtccctcagctgagg 216
Db 287 CCCAGGCCACACGCTGGCGGCTGGCTGAGGGAAACATGGCTGTGGCCCTCAGCTGAGG 346
QY 217 ttgctgctgtggaagaacctcactttcagagaagacaaacaglaagcttggttttca 276
Db 347 TTGCTGCTGTGGGAAGAACCTCACCTTTTCAGAGAAGACAAACATGTCAGCTGTACTGGAA 406
QY 277 gcagcggggggtctctcatt 321
Db 407 GTGGCCTGGCCTCTATTATCTTCTGATCTGATCTGATCTGCTGCTGCTGCTGCTGCTG 451

RESULT 14
AX060721
LOCUS AX060721 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078972.
ACCESSION AX060721
VERSION AX060721.1 GI:12406109
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 10474)
AUTHORS Lawn,R.M., Wade,D. and Garvin,M.
TITLE Regulation with binding cassette transporter protein abcl
JOURNAL Patent: WO 0078972-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN

Query Match 44.6%; Score 159.4; DB 6; Length 10474;
Best Local Similarity 81.8%; Pred. No. 1.2e-37;
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagtttaataaccagccacggcgctccctgtcagctgtgcccgtgcttcacagggct 156
Db 227 CAGTTTAATGACCAAGCCACGGCGCTCCCTGCTGTGAGCTGTGGCCGCTGCCAGGGCT 286
QY 157 cccagaccacacgctggcgtgctgctgaggaacatggcgtgttggtcagctgagg 216
Db 287 CCCAGGCCACACGCTGGCGGCTGGCTGAGGGAAACATGGCTGTGGCCCTCAGCTGAGG 346
QY 217 ttgctgctgtggaagaacctcactttcagagaagacaaacagtaagcttggttttca 276
Db 347 TTGCTGCTGTGGGAAGAACCTCACCTTTTCAGAGAAGACAAACATGTCAGCTGTACTGGAA 406
QY 277 gcagcggggggtctctcatt 321
Db 407 GTGGCCTGGCCTCTATTATCTTCTGATCTGATCTGATCTGCTGCTGCTGCTGCTGCTG 451

RESULT 15
AX060898
LOCUS AX060898 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 7 from Patent WO0078971.
ACCESSION AX060898
VERSION AX060898.1 GI:12406275
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 10474)
AUTHORS Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE Atp binding cassette transporter protein abcl polypeptides

JOURNAL Patent: WO 0078971-A 7 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others
ORIGIN

Query Match 44.6%; Score 159.4; DB 6; Length 10474;
Best Local Similarity 81.8%; Pred. No. 1.2e-37;
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagtttaataaccagccacggcgctccctgtcagctgtgcccgtgcttcacagggct 156
Db 227 CAGTTTAATGACCAAGCCACGGCGCTCCCTGCTGTGAGCTGTGGCCGCTGCCAGGGCT 286
QY 157 cccagaccacacgctggcgtgctgctgaggaacatggcgtgttggtcagctgagg 216
Db 287 CCCAGGCCACACGCTGGCGGCTGGCTGAGGGAAACATGGCTGTGGCCCTCAGCTGAGG 346
QY 217 ttgctgctgtggaagaacctcactttcagagaagacaaacagtaagcttggttttca 276
Db 347 TTGCTGCTGTGGGAAGAACCTCACCTTTTCAGAGAAGACAAACATGTCAGCTGTACTGGAA 406
QY 277 gcagcggggggtctctcatt 321
Db 407 GTGGCCTGGCCTCTATTATCTTCTGATCTGATCTGATCTGCTGCTGCTGCTGCTGCTG 451

Search completed: September 20, 2002, 03:13:43
Job time: 13937 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:08:22 ; Search time 520.94 Seconds
(without alignments)
1176.601 Million cell updates/sec

Title: US-09-846-456-2
Perfect score: 357
Sequence: 1 tggaggtctcagctgaggg.....gagggaaggaagctgtgttg 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	351.2	98.4	183999	22	AAF92831 Human ABC1 genomic
2	340.2	95.3	10545	21	AAC69132 Human ABC1 gene ex
3	159.4	44.6	10442	22	AAF24680 Nucleotide sequenc
4	159.4	44.6	10442	22	AAF24702 Nucleotide sequenc
5	159.4	44.6	10474	22	AAF24685 Nucleotide sequenc
6	159.4	44.6	10474	22	AAF24686 Nucleotide sequenc
7	159.4	44.6	10474	22	AAF24707 Nucleotide sequenc
8	159.4	44.6	10474	22	AAF24708 Nucleotide sequenc
9	145.8	40.8	446	22	AA504035 Partial human ABC1

10	145.8	40.8	7086	22	ABA09200 Human ABCA1 homolo
11	145.8	40.8	7086	22	AAK52667 Human polynucleoti
12	145.8	40.8	7260	22	AAJ21326 Human ATP binding
13	145.8	40.8	7260	22	AAJ70315 Human ATP binding
14	145.8	40.8	7281	22	AAK51683 Human polynucleoti
15	145.8	40.8	9741	22	AA506120 Human ABC1 DNA seq
16	145.8	40.8	9854	22	AAS06121 Human ABC1 DNA seq
17	144.2	40.4	736	22	AAH07432 Human ABC1 clone (
18	144.2	40.4	1556	22	AAH18606 Human cDNA sequenc
19	135.8	38.0	7857	21	AAC69388 Human ABC1 cholet
20	135.8	38.0	7860	22	AAF83826 Human ABC1 cholet
21	135.8	38.0	7860	22	AAF92835 Human ABC1 cDNA.
22	135.8	38.0	7861	21	AAC69387 Human ABC1 cholet
23	135.8	38.0	7864	21	AAC69120 Human ABC1 cholet
24	135.8	38.0	7864	21	AAC69385 Human ABC1 cholet
25	135.8	38.0	7864	21	AAC69386 Human ABC1 cholet
26	135.8	38.0	7864	21	AAC69389 Human ABC1 cholet
27	42.6	11.9	23024	22	AAF25499 Nucleotide sequenc
28	36.4	10.2	38	22	AAF93082 ABC1 polymorphism
29	36	10.1	7661	22	AAI6545 Human novel protei
30	36	10.1	7661	22	AAI62954 Human genomic DNA
31	36	10.1	7661	22	AAI63983 Human polynucleoti
32	36	10.1	11696	22	AAI6546 Human novel protei
33	36	10.1	11696	22	AAI6547 Human novel protei
34	36	10.1	11696	22	AAK72951 Human immune/haema
35	36	10.1	11696	22	AAK72952 Human immune/haema
36	36	10.1	11696	22	AAI62955 Human genomic DNA
37	36	10.1	11696	22	AAI62956 Human genomic DNA
38	36	10.1	11696	22	AAI63984 Human polynucleoti
39	36	10.1	11696	22	AAI63985 Human polynucleoti
40	36	10.1	11696	22	AAS33467 DNA encoding human
41	36	10.1	11696	22	AAS33468 DNA encoding human
42	35.8	10.0	534	21	AAK75311 Human ORFX ORF866
43	35.8	10.0	1133	22	AAD05589 Human secreted pro
44	35.8	10.0	50885	22	AAK70336 Human immune/haema
45	35.8	10.0	92407	22	AAF28549 Genomic fragment #

ALIGNMENTS

RESULT 1
AAF92831
ID AAF92831 standard; DNA; 183999 BP.
XX AAF92831;
AC AAF92831;
XX
DT 17-MAY-2001 (first entry)
XX Human ABC1 genomic DNA.
XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX Homo sapiens.
XX
XX WO200115676-A2.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-IB01492.
XX
XX 01-SEP-1999; 99US-0151977.
XX 15-MAR-2000; 2000US-0526193.
XX 23-JUN-2000; 2000US-0213958.
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX (XENO-) XENON GENETICS INC.
XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX WPI; 2001-244356/25.
XX
XX Treating a lower than normal high density lipoprotein-cholesterol

PT (HDL-C) level, a higher than normal triglyceride level, or a
 PT cardiovascular disease, by administering a compound that modulates LXR-
 PT or RXR-mediated transcriptional activity -

XX Claim 8; Fig 1; 317pp; English.

PS The present invention relates to a method for treating a patient
 CC diagnosed as having a lower than normal high density
 CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
 CC triglyceride level, or a cardiovascular disease, involving
 CC administering a compound that modulates LXR- or RXR-mediated
 CC transcriptional activity or ABC1 expression or activity.
 CC The LXR gene product may be used in an assay to identify
 CC compounds useful for the treatment of a disease or condition selected a
 CC lower than normal HDL cholesterol level, a higher than normal
 CC triglyceride level, and a cardiovascular disease.

XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 98.4%; Score 351.2; DB 22; Length 183999;
 Best Local Similarity 98.9%; Pred. No. 1.3e-97;
 Matches 353; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 tggaggtctcagctgagggctggtgattagcagtcctcattgtgtatggtttgcagca 60
 Db 53130 tggaggtctcagctgagggctggtgattagcagtcctcattgtgtatggtttgcagca 53189
 Qy 61 ataactgatggctgttccctcctcctgtctttatcttcaagtaataaccagcagcggt 120
 Db 53190 ataactgatggctgttccctcctcctgtctttatcttcaagtaataaccagcagcggt 53249
 Qy 121 cctctgtcagctctgagcgtgctcctccagggctccagccacacgctggcgctgct 180
 Db 53250 cctctgtcagctctgagcgtgctcctccagggctccagccacacgctggcggtgct 53309
 Qy 181 ggctgagggaacatggcatgttggcctcagctgaggtgtgctgtggaagacccctca 240
 Db 53310 ggctgagggaacatggctgttggcctcagctgaggtgtgctgtggaagacccctca 53369
 Qy 241 ttcagagaagacaaacagtaagcttggtttttcagcagcgggggttctcatttt 300
 Db 53370 ttcagagaagacaaacagtaagcttggtttttcagcagcgggggttctcatttt 53429
 Qy 301 tcttgtggttttggattgggattggagggagggagggagggaggaagctgtgtg 357
 Db 53430 tcttgtggttttggattgggattggagggagggagggaggaagctgtgtg 53486

RESULT 2

AAC69132
 ID AAC69132 standard; DNA; 10545 BP.

XX AAC69132;

XX 29-JAN-2001 (first entry)

DE Human ABC1 gene exon 1 (promoter).

KW Human ABC1 cholesterol transporter; chromosome 9q31; promoter;
 KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
 KW Tangier disease; Td; familial HDL deficiency; FHA; polymorphism;
 KW cardiovascular disease; coronary artery disease; coronary restenosis;
 KW cerebrovascular disease; peripheral vascular disease;
 KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
 KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
 KW prognosis; prophylaxis; drug screening; transgenic animal; ss.

OS Homo sapiens.

XX WO200055318-A2.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-IB00532.
 XX 15-MAR-1999; 99US-0124702.
 PR 08-JUN-1999; 99US-0138048.
 PR 17-JUN-1999; 99US-0139600.
 PR 01-SEP-1999; 99US-0151977.
 XX (UVR-) UNIV BRITISH COLUMBIA.
 PA (XENO-) XENON BIORESEARCH INC.
 PA Hayden MR, Wilson AR, Pimstone SN;
 PI WPI; 2000-587528/55.
 DR New ABC1 polypeptide is useful for treating diseases associated with
 PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
 PT disease and cancer -

XX Claim 50; Fig 12; 229pp; English.

XX The invention relates to the human ABC1 cholesterol transporter protein
 CC (B3082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of
 CC proteins, and plays a crucial role in cholesterol transport, particularly
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
 CC located on chromosome 9q31, and mutations in this gene are associated
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
 CC are distinguishable in that TD is an autosomal recessive disorder, while
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
 CC cholesterol") in the blood correlate with a high risk of cardiovascular
 CC disease, particularly coronary artery disease, but also cerebrovascular
 CC disease, coronary restenosis, and peripheral vascular disease.
 CC Conversely, a high level of HDL has protective effects against
 CC cardiovascular disease. The invention provides genetic constructs and
 CC transgenic cells and non-human animals comprising human ABC1 nucleic
 CC acids, and methods of gene therapy for the treatment or prevention of
 CC cardiovascular disease comprising the administration of an expression
 CC vector encoding ABC1 or an active fragment thereof. The invention also
 CC encompasses compounds which mimic ABC1 activity, compounds which
 CC stimulate ABC1 expression and methods of screening for such compounds.
 CC It further relates to methods for determining whether a patient has an
 CC increased risk for cardiovascular disease due to polymorphisms in the
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
 CC or prevent cardiovascular disease, especially coronary artery disease,
 CC cerebrovascular disease, coronary restenosis or peripheral vascular
 CC disease. They may also be used in the treatment of diseases associated
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
 CC The invention specifically excludes proteins with the exact amino acid
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
 CC present sequence represents the human ABC1 gene promoter region (exon 1).

XX Sequence 10545 BP; 2647 A; 2225 C; 2411 G; 3256 T; 6 other;

Query Match 95.3%; Score 340.2; DB 21; Length 10545;
 Best Local Similarity 98.9%; Pred. No. 9.5e-95;
 Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 tggaggtctcagctgagggctggtgattagcagtcctcattgtgtatggtttgcagca 60
 Db 8043 tggaggtctcagctgagggctggtgattagcagtcctcattgtgtatggtttgcagca 8102
 Qy 61 ataactgatggctgttccctcctcctgtctttatcttcaagtaataaccagcagcggt 120
 Db 8103 ataactgatggctgttccctcctcctgtctttatcttcaagtaataaccagcagcggt 8161
 Qy 121 cctctgtcagctctgagcgtgctcctccagggctccagccacacgctggcggtgct 180


```

FT FT      /*tag= a
XX FT      /product= "defective ABC1 polypeptide"
XX PN      WO200078972-A2.
XX PD      28-DEC-2000.
XX PF      16-JUN-2000; 2000WO-US16765.
XX PR      18-JUN-1999; 99US-0140264.
XX PR      14-SEP-1999; 99US-0153872.
XX PR      19-NOV-1999; 99US-0166573.
XX PA      (CVTH-) CV THERAPEUTICS INC.
XX PI      Lawn RM, Wade D, Garvin M;
XX PI      WPI; 2001-137812/14.
XX DR      Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
XX FT      useful for the development of agents for the treatment of heart disease
XX PT      and other disorders associated with hypercholesterolemia and
XX PT      atherosclerosis.
XX PS      Disclosure; Page 170-176; 215pp; English.
XX CC      The present sequence encodes a human adenosine triphosphate (ATP)
XX CC      binding cassette protein (ABC) 1 polypeptide, and is isolated from
XX CC      a Tangier disease patient. ABC1 resides in cell membranes and utilises
XX CC      ATP hydrolysis to transport a wide variety of substrates across the
XX CC      plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
XX CC      mobilisation of intracellular cholesterol stores. ABC1 is defective in
XX CC      Tangier disease, a genetic disorder characterised by abnormal
XX CC      HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
XX CC      9q22-q31. The ABC1 genes and proteins are useful for developing
XX CC      pharmaceutical agents for the treatment of heart disease and other
XX CC      disorders associated with hypercholesterolemia and atherosclerosis. The
XX CC      genes are useful for developing screening assays to screen for compounds
XX CC      that regulate the expression of genes associated with cholesterol
XX CC      transport. The genes and proteins are also useful for are also useful
XX CC      as diagnostic indicators of cardiovascular disease and other disorders
XX CC      associated with hypercholesterolemia.
XX SQ      Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

Query Match      44.6%; Score 159.4; DB 22; Length 10474;
Best Local Similarity 81.8%; Pred. No. 9.3e-39;
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagttaatgaccagcagcggtccctgctgtcagctgtgcccgtgcctccagggt 156
   |||||
Db 227 cagttaatgaccagcagcggtccctgctgtgagctgtgcccgtgcctccagggt 286
   |||||
QY 157 cccagaccacgctggcgctgctggtgaggaacatggcatgttgccctcagctgagg 216
   |||||
Db 287 cccagaccacgctggcgctgctggtgaggaacatggctgttgccctcagctgagg 346
   |||||
QY 217 ttgctgtgtggaagaacctcacttcacaggaagacaaacagtaagcttggttttca 276
   |||||
Db 347 ttgctgtgtggaagaacctcacttcacaggaagacaaacatgctgactgttactggaa 406
   |||||
QY 277 gcagcggggggttctctcattttttctttgtgtgttggttggtgg 321
   |||||
Db 407 gtggcctggcctctatttctctcctgactgactgactctgttctgg 451
   |||||

RESULT 7
AAF24707
ID AAF24707 standard; DNA; 10474 BP.
XX AC
XX AC
XX AAF24707;
XX

```

```

DT 20-APR-2001 (first entry)
XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
DE
XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 323..7108
XX FT /*tag= a
XX FT /product= "defective ABC1 polypeptide"
XX PN WO200078971-A2.
XX PD 28-DEC-2000.
XX PF 16-JUN-2000; 2000WO-US16591.
XX PR 18-JUN-1999; 99US-0140264.
XX PR 14-SEP-1999; 99US-0153872.
XX PR 19-NOV-1999; 99US-0166573.
XX PA (CVTH-) CV THERAPEUTICS INC.
XX PA (UNIW ) UNIV WASHINGTON.
XX PI Lawn RM, Wade D, Ozam JF, Garvin M;
XX PI WPI; 2001-137811/14.
XX DR P-PSDB; AAB31366.
XX CC Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX CC polynucleotides and polypeptides, useful for treatment of heart disease
XX CC and other disorders associated with hypercholesterolemia and
XX CC atherosclerosis.
XX PS Claim 27; Page 144-150; 211pp; English.
XX CC The present sequence encodes a human adenosine triphosphate (ATP)
XX CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
XX CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
XX CC ATP hydrolysis to transport a wide variety of substrates across the
XX CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
XX CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
XX CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
XX CC 9q22-q31. The ABC1 genes and proteins are useful for developing
XX CC pharmaceutical agents for the treatment of heart disease and other
XX CC disorders associated with hypercholesterolemia and atherosclerosis. The
XX CC genes are useful for developing screening assays to screen for compounds
XX CC that regulate the expression of genes associated with cholesterol
XX CC transport. The genes and proteins are also useful for are also useful
XX CC as diagnostic indicators of cardiovascular disease and other disorders
XX CC associated with hypercholesterolemia.
XX SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match      44.6%; Score 159.4; DB 22; Length 10474;
Best Local Similarity 81.8%; Pred. No. 9.3e-39;
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagttaatgaccagcagcggtccctgctgtcagctgtgcccgtgcctccagggt 156
   |||||
Db 227 cagttaatgaccagcagcggtccctgctgtgagctgtgcccgtgcctccagggt 286
   |||||
QY 157 cccagaccacgctggcgctgctggtgaggaacatggcatgttgccctcagctgagg 216
   |||||
Db 287 cccagaccacgctggcgctgctggtgaggaacatggctgttgccctcagctgagg 346
   |||||

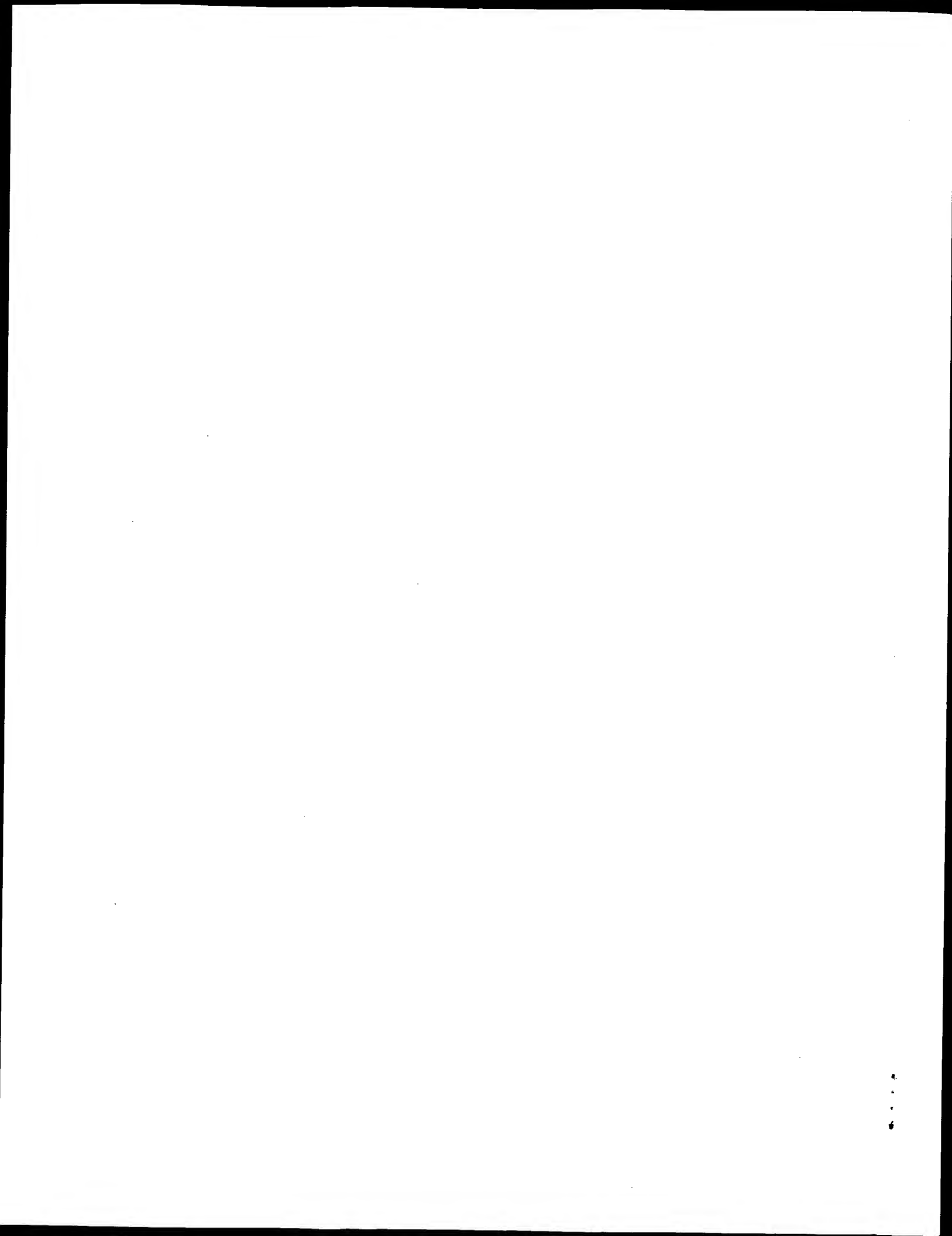
```


XX New adenosine triphosphate binding cassette transporter gene
PT polymorphisms, useful for diagnosing and treating lipid disorders,
PT cardiovascular diseases and inflammatory diseases -
XX
XX Disclosure: Page 26-28; 41pp; English.
XX
XX The present sequence is that of cDNA encoding the human adenosine
CC triphosphate (ATP) binding cassette transporter 1 (ABCL1) protein
CC (see AAM50227). The sequence includes an extended open reading
CC frame (ORF) to that provided by the sequence in AAI70314, using
CC an alternative ATG codon as initiation codon and thereby adding an
CC

PR 19-JUL-2000; 2000US-0620325.

us-09-846-456-2.std.rng

Fri Sep 20 08:03:38 2002



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:21:22 ; Search time 110.16 Seconds
(without alignments)
796.035 Million cell updates/sec

Title: US-09-846-456-2

Perfect score: 357

Sequence: 1 tggaggtctcagctgagagg.....gagggaaggagctgtgttg 357

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA:*
- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	8.7	10968	2	US-08-680-327-2
2	31	8.7	10968	4	US-09-228-246-1
3	30.8	8.6	1037	1	US-08-462-195-3
4	30.8	8.6	1037	2	US-08-636-883-3
5	30.8	8.6	1037	3	US-09-127-829-3
6	30.4	8.5	2473	4	US-09-173-914-3
7	30	8.4	686	4	US-08-988-321B-37
8	30	8.4	686	4	US-08-397-220B-25
9	30	8.4	4771	2	US-08-866-650-2
10	30	8.4	4771	2	US-09-021-287-2
11	30	8.4	4771	4	US-09-240-473-2
12	30	8.4	5143	1	US-08-574-043A-7
13	30	8.4	5143	2	US-08-795-015-7
14	30	8.4	7881	2	US-08-751-189-1
15	30	8.4	7881	2	US-09-060-836-1
16	30	8.4	7881	4	US-09-184-445-1
17	30	8.4	3529	4	US-09-144-085-3
18	29.6	8.3	424	1	US-08-609-637-15
19	29.4	8.2	2480	4	US-09-534-638-3
20	29.4	8.2	3147	4	US-08-781-802-7
21	29.4	8.2	3147	4	US-08-694-078-7
22	29.4	8.2	3147	4	US-09-058-260-7
23	29.4	8.2	31491	4	US-09-360-186-1
24	29.2	8.2	573	2	US-08-290-665A-124
25	29.2	8.2	573	5	PCR-US95-10398-124
26	29.2	8.2	2892	1	US-08-264-534-5
27	29.2	8.2	2892	1	US-08-083-590A-1

c 28	29.2	8.2	2892	1	US-08-465-500-5	Sequence 5, Appli
c 29	29.2	8.2	2892	2	US-08-346-126-5	Sequence 5, Appli
c 30	29.2	8.2	2892	2	US-08-346-128-5	Sequence 5, Appli
c 31	29.2	8.2	2892	3	US-08-532-384-1	Sequence 1, Appli
c 32	29.2	8.2	2892	3	US-08-893-828-5	Sequence 5, Appli
c 33	28.8	8.1	1257	3	US-08-640-906-1	Sequence 1, Appli
c 34	28.8	8.1	1257	4	US-09-395-936-1	Sequence 10, Appli
c 35	28.6	8.0	2861	4	US-08-482-073-10	Sequence 7, Appli
c 36	28.6	8.0	3647	1	US-07-914-281-7	Sequence 7, Appli
c 37	28.6	8.0	3647	1	US-08-393-246-7	Sequence 7, Appli
c 38	28.6	8.0	3647	1	US-08-525-058A-7	Sequence 7, Appli
c 39	28.6	8.0	3647	2	US-08-696-731-7	Sequence 7, Appli
c 40	28.6	8.0	3647	4	US-09-042-531-7	Sequence 7, Appli
c 41	28.4	8.0	345	1	US-08-324-977-7	Sequence 7, Appli
c 42	28.4	8.0	345	2	US-08-384-616-7	Sequence 7, Appli
c 43	28.4	8.0	345	2	US-08-904-886A-7	Sequence 7, Appli
c 44	28.4	8.0	345	4	US-09-315-850-7	Sequence 7, Appli
c 45	28.4	8.0	573	2	US-08-290-665A-115	Sequence 115, App

ALIGNMENTS

RESULT 1
US-08-680-327-2
; Sequence 2, Application US/08680327
; Patent No. 5859321
; GENERAL INFORMATION:
; APPLICANT: Staskawicz, Brian S., Oldroyd, Giles Edward,
; APPLICANT: Salmeron, John M., Rommens, Caius
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT
; TITLE OF INVENTION: PATHOGEN RESISTANCE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &
; ADDRESSEE: Whinston
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street
; STREET: Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk, 3-1/2 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,327
; FILING DATE: July 11, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/310,912
; FILING DATE: September 22, 1994
; CLASSIFICATION: 800
; APPLICATION NUMBER: 08/227,360
; FILING DATE: April 13, 1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Alan. E.
; REGISTRATION NUMBER: 35,123
; REFERENCE/DOCKET NUMBER: 5151-45038
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10968 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
US-08-680-327-2

; Sequence 3, Application US/08636883
; Patent No. 5830691
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIHARU
; APPLICANT: HONDA, YOSHIKAZU
; APPLICANT: SEKI, MAKOTO
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/636,883
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE:
; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5830691man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1037 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: pUCM010
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..1036
US-08-636-883-3

Query Match 8.6%; Score 30.8; DB 2; Length 1037;
Best Local Similarity 46.3%; Pred. No. 2.1; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 117

QY 107 ccagccacggcgctccctgtcagctctggccgctgctccagggctccagcgac 166
DB 237 CTACCCCTCGGTTGGCGAGCCTTGGGATAGTTGTCGCTTCCACGAGGTTCGCGCGC 178
QY 167 acgctggcgctgctggctgaggaacatggcatgttgccctcagctgaggttctgt 226
DB 177 TCGGAAGTCTTCTTAGTTCGCGCGCACACCAACCTGGGGCCCTTCGCGCGCAACAGGTAA 118

QY 227 ggaagaacctcacttcagaaagaagacaaacagtaagcttgggtttttcagcagcgggg 286
DB 117 ACTCCACCAACGATCTGACACCGCCCGGGAACCTTACGTCCTGTGGCGGCGGTGGTG 58
QY 287 gtctctcattttttttttgtgtgtttttgtgattggggat 324
DB 57 TTACGTTTGATTTTCTTTGGGGTTTGGGATTGTGCT 20
RESULT 5
US-09-127-829-3/c
; Sequence 3, Application US/09127829
; Patent No. 6063904
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIHARU
; APPLICANT: HONDA, YOSHIKAZU
; APPLICANT: SEKI, MAKOTO
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,829
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE:
; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 6063904man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1037 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: pUCM010
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..1036
US-09-127-829-3

Query Match 8.6%; Score 30.8; DB 3; Length 1037;

```
; Best Local Similarity 46.3%; Pred. No. 2.1;
Matches 101; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 107 ccagccacggggtcctctgctcagctgtgcccctgctccaggggtcccgagccac 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 CTACCCCTGGGTGGGAGCCCTGGGAGTAGGTTCGCCCTCCACGAGGTTCGCCGCGC 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 167 acgctggggtgctgctgaggaacatggaatgttgccctcagctgaggtgtgctgt 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 TCGGAAGTCTTCTAGTCGCGGCACACCCAACTGGGGCCCTCGCGGCAACAGGTAA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 ggaagaacctcacttcagaagaagaaacagtaagcttggtttttcagcagcggggg 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 ACTCCACCAACGATGACACCGCGCGGAACCTTAACGTCCTGTGGCGCGGTGGTG 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 gttctctcatttttttttttttttttttttttttttttttttttttttttttt 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 TTACGTTTGATTTTCTTTGGGGTTTGGATTGTGCT 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-173-914-3/c
; Sequence 3, Application US/09173914
; Patent No. 6171857
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Eric
; TITLE OF INVENTION: A No. 6171857el Leucine Zipper, KARP-1 and
; FILE REFERENCE: B0877/7017/HK
; CURRENT APPLICATION NUMBER: US/09/173,914
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 60/064,557
; EARLIER FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Mesocricetus Auratus
US-09-173-914-3

Query Match 8.5%; Score 30.4; DB 4; Length 2473;
Best Local Similarity 57.3%; Pred. No. 4.5;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 250 agacaaacagtagctgtggtttttcagcagcgggggttctctcattttttttgtg 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1629 AGAGGAATCTACATTTGGCTTTTCCCTCAGTGGGGGATCCAGCATATTCAAAATGTGC 1570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 ttttgattggggattggaggaggaggaggaggaggaggaggaggaggaggaggagg 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1569 TGCTGAATTTGGGGTAGAGCCTCTCTGGGGATGTAAG 1534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-08-988-321B-37/c
; Sequence 37, Application US/08988321B
; Patent No. 6174868
; GENERAL INFORMATION:
; APPLICANT: Kevin P. Anderson et al.
; TITLE OF INVENTION: Compositions And Methods For Treatment Of Hepatitis C V
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Jane Massey Licata
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM COMPATIBLE
```

```
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.1 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,321B
; FILING DATE: December 10, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/650,093
; FILING DATE: May 17, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/452,841
; FILING DATE: May 30, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/397,220
; FILING DATE: March 9, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/945,289
; FILING DATE: September 10, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0245
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686
; TYPE: nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
US-08-988-321B-37

Query Match 8.4%; Score 30; DB 4; Length 686;
Best Local Similarity 46.3%; Pred. No. 3.1;
Matches 99; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 111 ccacgggggtcctgtgtcagctctgtgcccgtgcttccaggggtcccgagccacacgc 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CCTCGGCGCGGAGCCTTGGGGATAGCTGTGCGCTTCCACGGGGTTGCGACCGCTCGG 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 tggcgctgctggtgagggaaacatg9catgttgccctcagctgaggtgtgctgtgtgaa 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 AAGTCTTCTCTGATCGCGCGCACACCCAACTGGGGCCCTCGCGGCAACAGGTAAATC 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 231 gaacctcaccttcagaagaagaaacagtaagcttgggttttttcagcagcgggggttc 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 CACCAACGATCTGACACCGCGCGGGAACCTTGACCTCTGTGGCGCGGTTGGTGTAC 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 tctcattttttttttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTTTGGTTTTCCTTTGAGGTTAGGATTCGTGCT 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-08-397-220B-25/c
; Sequence 25, Application US/08397220B
; Patent No. 6284458
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: Compositions And Methods For Treatment Of Hepatitis C Virus-Associated Diseases
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
```



```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mTll protein"
; US-09-021-287-2

Query Match      8.4%; Score 30; DB 2; Length 4771;
Best Local Similarity 52.4%; Pred. No. 8.5;
Matches 66; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 102 aatgaccagcagcggcgctccctgtcgtcagctctgcccgcgtgccttccagggtcccca 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 ATTCACTTCCCGGAGCAGCGCGGTGGCAGCGCGCGGCGCGGCGCGCTGCAGCTCGCTC 200

QY 162 gccacacgctggcgctgctgagtgaggaacatggcatgttgcctcagctgaggttgct 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 201 TCGGCGCGGGGTCTTACAGCGCGCGGGCGGCGCGGCGCGGAGCTCCGCTGGCA 260

QY 222 gctgtg 227
    |||||
Db 261 GCTGAG 266

RESULT 11
US-09-240-473-2
; Sequence 2, Application US/09240473
; Patent No. 6297011
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hofman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/240,473
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mTll protein"
; US-09-240-473-2

Query Match      8.4%; Score 30; DB 4; Length 4771;
Best Local Similarity 52.4%; Pred. No. 8.5;
Matches 66; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 102 aatgaccagcagcggcgctccctgtcgtcagctctgcccgcgtgccttccagggtcccca 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 ATTCACTTCCCGGAGCAGCGCGGTGGCAGCGCGCGGCGCGGCGCTGCAGCTCGCTC 200

QY 162 gccacacgctggcgctgctgagtgaggaacatggcatgttgcctcagctgaggttgct 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 201 TCGGCGCGGGGTCTTACAGCGCGCGGGCGGCGCGGCGGAGCTCCGCTGGCA 260

QY 222 gctgtg 227
    |||||
Db 261 GCTGAG 266

RESULT 12
US-08-574-043A-7/c
; Sequence 7, Application US/08574043A
; Patent No. 5807692
; GENERAL INFORMATION:
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: El-Deiry, Wafik
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: p21WAF1 Derivatives and Diagnostic
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW suite 1100
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,043A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.49698
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

Search completed: September 20, 2002, 03:21:31
Job time: 14180 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:10 : Search time 3895 Seconds
(without alignments)
1237.076 Million cell updates/sec

Title: US-09-846-456-2
Perfect score: 357
Sequence: 1 tggaggtctcagctgaggg.....gagggaaggagctgtgttg 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_estl: *
10: gb_est2: *
11: gb_htc: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vrt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144.2	40.4	736	9	AU135588
2	75.8	21.2	619	9	BB657864
3	73.8	20.7	292	10	244377 HSC1ZB081 n
4	71.2	19.9	535	10	BG384217 303216 MA
5	42.4	11.9	982	12	AL192682 Tetraodon
6	39.6	11.1	910	12	CNS02D01
7	37.4	10.5	926	12	CNS0060N
8	37	10.4	440	9	AZ542175
9	37	10.4	533	9	AA914462
10	36.4	10.2	571	9	AV845237
11	36.4	10.2	579	9	AV864011
12	36	10.1	322	9	AV892280
13	36	10.1	393	9	AU056364
14	36	10.1	571	9	AU162694
15	36	10.1	872	12	AV862022
16	35.8	10.0	344	10	AL184552 Tetraodon
17	35.8	10.0	785	12	BM149133 TCAAP2E63
					AF010859 AF010859

C	18	35.8	10.0	939	12	CNS00CNG
	19	35.6	10.0	514	12	BH087817
C	20	35.6	10.0	1097	12	CNS04PHC
	21	35.6	10.0	1574	10	BG325753
	22	35.4	9.9	390	9	BB841907
	23	35.4	9.9	482	12	AQ484626
C	24	35.4	9.9	937	12	CNS006ST
C	25	35.4	9.9	1901	10	BF128237
C	26	35.2	9.9	972	10	BE733920
	27	35.2	9.9	1201	12	CNS0162E
C	28	34.8	9.7	301	9	AW751638
	29	34.8	9.7	365	9	BB843149
	30	34.8	9.7	516	10	BF484412
	31	34.8	9.7	530	10	BG471178
C	32	34.8	9.7	810	10	BE748387
	33	34.8	9.7	1068	12	CNS00ETV
	34	34.6	9.7	483	12	CNS03210
	35	34.6	9.7	884	12	CNS006U0
	36	34.6	9.7	1039	12	CNS015P6
C	37	34.6	9.7	1068	12	CNS00ETV
	38	34.6	9.7	1101	12	CNS00LT2
	39	34.4	9.6	460	9	AU212427
C	40	34.4	9.6	778	10	BI886434
C	41	34.2	9.6	619	9	AW681359
	42	34.2	9.6	896	12	CNS00BP8
	43	34.2	9.6	970	12	CNS010C9
C	44	34.2	9.6	1201	12	CNS01604
C	45	34	9.5	468	12	AQ773862

ALIGNMENTS

RESULT 1

AU135588
LOCUS
DEFINITION AU135588 PLAC1 Homo sapiens cDNA clone PLACE1002437 5', mRNA linear EST 24-OCT-2000
sequence.
ACCESSION AU135588.1 GI:10996127
VERSION AU135588
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 736)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
source
Location/Qualifiers
1..736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"
/clone_lib="PLACE1"
/tissue_type="placenta"
/notes="Vector: pME18SFL3"
BASE COUNT 163 a 199 c 199 g 170 t 5 others
ORIGIN

Query Match 40.4%; Score 144.2; DB 9; Length 736;
 Best Local Similarity 80.4%; Pred. No. 8.3e-28;
 Matches 181; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 97 cagtaataaccagcagcggcctccctgctcaagctcgtgcccgtgctccagggct 156
 |||||
 Db 219 CAGTTAATACCAACCCAC-GGGCTCCCTGCTGTGAGCTGTGGCCGCTTCCAGGGCT 277
 |||||
 QY 157 cccgagcacacagcctggcgtgctgaggaacatggcattgttggcctcagctcagctgag 216
 |||||
 Db 278 CCCGAGCACACGCTGGGGGTGCTGCTGAGGGAACATGCTGTGTGGCCCGCAGCTGAGG 337
 |||||
 QY 217 ttgctgctgtggaagaacctcaatttcagaagaagacaagtaagcttgggtttttca 276
 |||||
 Db 338 TTGCTGCTGTGGAAGAACCTCACTTTTCAAGAAGACAACATGTGAGCTGTGCTGGAA 397
 |||||
 QY 277 gcagcggggggtctctcatttttttttttgggttttggattggg 321
 |||||
 Db 398 GTGGCTGGCTCTATTATCTTCCTGATCCTGATCTGTTCGG 442

RESULT 2

BB657864
 LOCUS BB657864 619 bp mRNA linear EST 26-OCT-2001
 DEFINITION musculus cDNA clone D230019D04 5', mRNA sequence.
 BB657864
 ACCESSION BB657864.1 GI:16491690
 VERSION EST.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 619)

REFERENCE

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/

TITLE

JOURNAL

COMMENT

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10
 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.

FEATURES

source

Location/Qualifiers
 1..619
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="D230019D04"
 /clone_lib="RIKEN full-length enriched, 12 days embryo
 eyeball"
 /tissue_type="eyeball"
 /dev_stage="12 days embryo"
 /lab_host="DH10B"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGCGCGCGCAACTCGAGTTTTTTTTTTTTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGAGATTCTCGAGTTAAATTAATCCGCCCCCCCCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda FLC I."
 BASE COUNT 126 a 185 c 173 g 134 t 1 others
 ORIGIN

Query Match 21.2%; Score 75.8; DB 9; Length 619;
 Best Local Similarity 77.8%; Pred. No. 1.1e-09;
 Matches 133; Conservative 0; Mismatches 27; Indels 11; Gaps 3;
 QY 97 cagtaataaccagcagcggcctccctgctcagctcgtgcccgtgctccagggct 156
 |||||
 Db 219 CAGTTAATGACCAAGCCAC-AGAGTCACAGCTCTGTGCTGTGGCTGCT-CCCTCCAGGGCT 276
 |||||
 QY 157 cccgagcacaca-----cgctggcgctgctgctgaggggaacatggcattgtggcct 207
 |||||
 Db 277 CTCGAGCGCGCAGAGCGAGCTGCTGTGGTGGCGGCTGTGTGACATGGCTTGTGGCCT 336
 |||||
 QY 208 cagctgagttgctgctggaagaaacactcatttcagaagaagacaaca 258
 |||||
 Db 337 CAGTTAAGGCTGCTGTGTGGAGAATCTGACATTCGAGGAGAGACAACA 387

RESULT 3

244377
 LOCUS 244377 292 bp mRNA linear EST 14-NOV-1994
 DEFINITION HSC12B081 normalized infant brain cDNA Homo sapiens cDNA clone
 C-12B081, mRNA sequence.
 ACCESSION 244377
 VERSION 244377.1 GI:573506
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 292)
 REFERENCE
 AUTHORS Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes
 M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F.,
 Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
 Sebastiani-Kabatchis, C. and Tessier, A.
 IMAGE: molecular integration of the analysis of the human genome
 and its expression
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 95277534
 COMMENT Contact: Genethon

Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C; Genexpress_sequence_id: y1c-lzb08
Seq primer: (-21)M13_universal.
Location/Qualifiers
1. .292
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-lzb08"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: Brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 50 a 87 c 96 g 56 t 3 others
ORIGIN

Query Match 20.7%; Score 73.8; DB 10; Length 292;
Best Local Similarity 94.5%; Pred. No. 3.4e-09;
Matches 86; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 97 cagtaataaccagccagggcgtccctgtgtcagctctgcccgtgccttcacagggt 156
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 203 CAGTTAATGANCACACAC-GCGGTCCTGTGTGAGCTGTGGCGCTGCCTTCAGGGT 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 157 cccagccacagctggcgtgctgctgag 187
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 CCCGAGCCACACGCTGGGGTGTGGCTGAG 292
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
BG384217
LOCUS BG384217 535 bp mRNA linear EST 12-MAR-2001
DEFINITION 303216 MARC lPIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG384217
VERSION BG384217.1 GI:13308689
KEYWORDS pig.
SOURCE Sus scrofa
ORGANISM Sus scrofa
REFERENCE 1 (bases 1 to 535)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. bases called and alt.trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATCACCAT
BACKWARD: GTTTCCAGTCACGACG

Plate: 90 row: G column: 13
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .535
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC lPIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT 121 a 159 c 136 g 119 t
ORIGIN

Query Match 19.9%; Score 71.2; DB 10; Length 535;
Best Local Similarity 75.9%; Pred. No. 1.8e-08;
Matches 132; Conservative 0; Mismatches 28; Indels 14; Gaps 3;

Qy 97 cagtaataaccagccagggcgtccctgtgtcagctctgcccgtgccttcacagggt 156
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 CAGTTAATGACCAGCCACAG--CGCCCCCTGTGTGAGCTGTGGCCACTTCTCCACGGGC 235
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 157 ccc-----gagccacacgctg-----ggcgtgctggctgagggaacatgcatgttg 204
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 236 TCCGAGCGGAGCCACACGAGCTGTGTGTGGTGTGGAGTAACATGCTTATTGG 295
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 205 cctcagctgaggtgtgtgtgtggaagacctcacttcagaaagagacaaca 258
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 296 ACTCAACTGAGGTACTGTGTGGAAAGACCTCACTTTCAGAAAGACAAACA 349
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
CNS02D01/C
LOCUS CNS02D01/C 982 bp DNA linear GSS 13-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
26019 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL192682
VERSION AL192682.1 GI:7830786
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 982)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 982)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 982)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source
Location/Qualifiers
1. .982
/organism="Tetraodon nigroviridis"

```

/db_xref="taxon:99883"
/clone="260p19"
/clone_lib="G"
/note="Genoscope sequence ID : COAG260CH10SP1-end :
PUC-ori"
BASE COUNT      246 a  248 c  224 g  247 t  17 others
ORIGIN

Query Match      11.9%; Score 42.4; DB 12; Length 982;
Best Local Similarity 55.0%; Pred. No. 0.86;
Matches 66; Conservative 8; Mismatches 46; Indels 0; Gaps 0;

QY  237 cactttcagaagaacacagtgagcttggttttttcagcagcggttggtctctcat 296
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  839 CACACCGACTAATRAAKAWAYTTTCAGACCGGCTTBTCTGSRGKGGGGTCTTTT 780
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  297 tttttcttggttttttgattggagtgaggagggagggaggaaggaagctgttt 356
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db  779 TTTTTKTTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

RESULT 6
CNS0060N
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION      AL065629.1 GI:4944698
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 910)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoeawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .910
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR14J21"
/note="end : T7"
BASE COUNT      202 a  63 c  112 g  198 t  335 others
ORIGIN

Query Match      11.1%; Score 39.6; DB 12; Length 910;
Best Local Similarity 50.8%; Pred. No. 4.7;
Matches 64; Conservative 11; Mismatches 51; Indels 0; Gaps 0;

QY  229 aagaacctcacttcagaagaacacagtaagcttggttttttcagcagcggttg 288

```

```

||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  369 AAAAAAAAAAATAAAAAAAAAAAAAAAAAAKARWGGKAGGGCTKAGCGCDTGTGTGTT 428
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  289 tctctatttttttttggttggttgattggagggagggagggaggaaggaa 348
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  429 TTTTTTTTTTTTTTTTTKTGKTKTTTTTTTTTTTGTGWWGGGGGGGGKGGGGGGGGG 488
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  349 gctgtg 354
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  489 GGGGG 494

RESULT 7
AZ542175
LOCUS
DEFINITION
ENTGP86TF Entamoeba histolytica Sheared DNA linear GSS 14-NOV-2000
genomic, DNA sequence.
ACCESSION      AZ542175
VERSION
KEYWORDS
SOURCE
ORGANISM
Entamoeba histolytica.
Entamoeba histolytica.
Eukaryota; Entamoebidae; Entamoeba.
1 (bases 1 to 926)
REFERENCE
AUTHORS
Loftus, B., Van Aken, S. and Fraser, C.
TITLE
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL
Unpublished (2000)
COMMENT
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 297.
Location/Qualifiers
1. .926
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: PHOS1; Site1: Bst I; Constructed at The
Institute for Genomic Research (IGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
BASE COUNT      304 a  29 c  219 g  374 t
ORIGIN

Query Match      10.5%; Score 37.4; DB 12; Length 926;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY  228 gaagaaccttcacttcagaagaacacagtaagcttggttttttcagcagcggttg 287
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  757 GAAAAAGGGAAGAAAAAGAAAAAGAAAAAGAAAGTTGTTTGTGGGGGGGTGGG 816
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  288 tctctatttttttttggttggttgattggagggagggagggaggaagg 346

```


[REDACTED]

Search completed: September 20, 2002, 01:23:14
Job time: 7374 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:13:43 ; Search time 5225.75 Seconds
(without alignments)
884.995 Million cell updates/sec

Title: US-09-846-456-4

Perfect score: 221

Sequence: 1 gtaattgcgcgcgagtgta.....aacacaaagtggaaacag 221

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_ov.*

5: gb_pat.*

6: gb_ph.*

7: gb_pl.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vl.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	221	100.0	221	6	AX351032	Sequence	
2	221	100.0	697	9	AF258627	Homo sapi	
3	221	100.0	1167	9	HSJ25201	Homo sapi	
4	221	100.0	1167	9	AF258623S1	Homo sapi	
5	221	100.0	3231	6	AX351029	Sequence	
6	221	100.0	7260	6	AX253452	Sequence	
C	7	221	100.0	96717	9	AL359182	Human DNA
8	221	100.0	149034	9	AF275948	Homo sapi	
9	221	100.0	175064	2	AC012230	Homo sapi	
10	221	100.0	201144	9	AF287262	Homo sapi	
11	220.6	99.8	183999	6	AX092589	Sequence	
12	220	99.5	69570	2	AC021246	Homo sapi	
13	219.4	99.3	1556	9	AK024328	Homo sapi	
14	219	99.1	1750	9	AK022254	Homo sapi	
C	15	214.8	97.2	90698	2	AC021345	Homo sapi
16	205	92.8	9854	6	AX127831	Sequence	
17	205	92.8	9854	6	AX139818	Sequence	
C	18	203.4	92.0	69570	2	AC021246	Homo sapi
19	197	89.1	10442	6	AX060713	Sequence	
20	197	89.1	10442	6	AX060892	Sequence	
21	197	89.1	10442	9	AF285167	Homo sapi	
22	190	86.0	10474	6	AX060719	Sequence	
23	190	86.0	10474	6	AX060721	Sequence	
24	190	86.0	10474	6	AX060898	Sequence	
25	190	86.0	10474	6	AX060900	Sequence	
26	92	41.6	446	6	AX127764	Sequence	
27	92	41.6	446	6	AX139751	Sequence	
28	92	41.6	9741	6	AX127830	Sequence	
29	92	41.6	9741	6	AX139817	Sequence	
30	92	41.6	9741	6	AX351038	Sequence	
31	91	41.2	1643	6	AX060715	Sequence	
32	91	41.2	1643	6	AX060894	Sequence	
33	67.4	30.5	278572	10	AF287263	Mus muscu	
34	39.2	17.7	54705	2	AC103155	Rattus no	
C	35	38.2	17.3	35122	9	AP000302	Homo sapi
36	38.2	17.3	100000	9	AP000046	Homo sapi	
C	37	38.2	17.3	100000	9	AP000114	Homo sapi
C	38	38.2	17.3	100000	9	AP000190	Homo sapi
C	39	38.2	17.3	187272	2	AC096044	Rattus no
C	40	38.2	17.3	340000	9	AP001717	Homo sapi
41	37.2	16.8	13613	2	AC093942	Rattus no	
42	36.6	16.6	204924	2	AC015616	Homo sapi	
43	35.6	16.1	2344	10	MMVGR6	Mus musculu	
C	44	35.6	16.1	2674	10	BC016252	Mus muscu
45	35.4	16.0	110000	2	LMFLCHR31_18	Continuation (19 0	

ALIGNMENTS

RESULT	1	AX351032	Sequence 4 from Patent WO0183746.	221 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX351032	Sequence 4 from Patent WO0183746.					
DEFINITION	AX351032	Sequence 4 from Patent WO0183746.					
ACCESSION	AX351032	Sequence 4 from Patent WO0183746.					
VERSION	AX351032.1	GI:18616388					
KEYWORDS	human.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	1 (sites)						
AUTHORS	Brewer B., Duverger N., Renaley, A. and Santamarina-Fojo, S.						
TITLE	Regulatory nucleic acid sequences of the abcl gene						
JOURNAL	Patent: WO 0183746-A 4 08-NOV-2001;						
FEATURES	Aventis Pharma S.A. (FR)						
SOURCE	Location/Qualifiers						
	1..221						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
BASE COUNT	44 a	62 c	73 g	42 t			
ORIGIN							

986

D	b	987	CCGGGCTCGGCAGGGCAGGGCGGGGAGCTCCGCCGACCAACAGAGCGCGGTTCATCAGGGC	1046
Q	y	121	gccttgctcctgttttccccgggttcgtgattctccctctcccgaagcgcttgtaa	180
D	b	1047	GCTTTGCTCCTTGTTTTTCCCCGGTTCGTGTTTCTCCCTTCTCCGAAGGCTTGTCAA	1106
Q	y	181	ggggtagagaaagacgacgcaaacacaaaaagtggaaaacag	221
D	b	1107	GGGCTAGGAGAAAGAGACGCAACAACAAAAAGTGGAACACAG	1147
R	E	S		
A	X	351029		
L	O	CUS		
D	E	F	I	N
A	C	C	E	S
V	E	R	S	I
K	E	Y	W	O
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O	U	R	C
S	O	R	C	E
O	R	G	A	N
R	E	F	E	R
A	U	T	H	O
T	I	T	L	E
J	O	U	R	N
F	E	A	T	E
B	A	S	E	
S	O			

VERSION	AF287262.1	GI:13876612	
KEYWORDS	human.		
SOURCE	Human sapiens		
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 201144)		
REFERENCE	Qiu, Y., Cavellier, L., Chiu, S., Yang, X., Rubin, E. and Cheng, J.F. Human and mouse abcal comparative sequencing and transgenesis studies revealing novel regulatory sequences		
AUTHORS	Genomics 73 (1), 66-76 (2001)		
TITLE	21251004		
JOURNAL	2 (bases 1 to 201144)		
MEDLINE	Qiu, Y., Cavellier, L., Chiu, S., Rubin, E. and Cheng, J.-F. Direct Submission		
REFERENCE	Submitted (13-JUL-2000) Genome Science Department, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley, CA 94720, USA		
AUTHORS	Location/Qualifiers		
TITLE	1. .201144		
JOURNAL	/organism="Homo sapiens"		
FEATURES	/db_xref="taxon:9606"		
SOURCE	join(33931..34151,58319..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,12329..123569,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150428,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170229,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..181457)		
	/gene="ABCA1"		exon
gene	/product="ATP-binding cassette 1 sub-family A member 1"		exon
exon	33931..181457		exon
exon	/gene="ABCA1"		exon
exon	33931..34151		exon
exon	/gene="ABCA1"		exon
exon	/number=1		exon
exon	58319..58476		exon
exon	/gene="ABCA1"		exon
CDS	/number=2		exon
	join(58411..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,12329..123569,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150428,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170229,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..181457)		exon
	/gene="ABCA1"		exon
	/codon_start=1		exon
	/product="ATP-binding cassette 1 sub-family A member 1"		exon
	/protein_id="AAK43526.1"		exon
	/db_xref="GI:13876613"		exon
	/translation="MACWPQLRLILWKNLFRRTQTCLLLEAVNPLFILLISVRLSYPPQEHCFFPNKMPASGTLFWVGIIICNANNPCFRPTGEPAGVGVGNKSLYARLFSDARLLLYLQKDTSMKMDRKVLRTLQKIKSSNLKQDLFDVDFNFTGSLYHNLSPKSTVDKMLRADVLILKVFLLQYQLHLTSLCNGSKSEEMIQLDQDEVSELGILP		exon

KEKLAARERLVRNMDILKPIILRTLNSTSPSPSKELAEATKTLHLHSLGTLAQLFSMR
SKSDMQEVMELTNVNSSSSSTQIYQAVSRIVCGHPGGGLKIKSLNMYEDNNKALF
CGNGTEDEATFYIDNATPPLNLMKNLESSPLSLIKWALKPLIKVLTDPYDPAT
ROVMAEVNKTFOELAVFDLEGMEEELSPKIWTIMENSQEMDLVRLMLDSRDNDPWE
QOLQDLTAQDIVAFLAKHPEDVQSSSSYITWREAFNETNQAIRITSRMECVNLN
KLEPIATEWLINKSMELLDERKFWAGIVFTGIPGSELPHRVKIKRIMIDNVERT
NKIKDGYWDPGRADPDMRYVNGFAYLDQVVEQAIIRVLTGTEKTKYVYMOOMPY
PCYDDIFPLRVSRMPLFMTLAWISYAVTIKIVYEKEARLEKTRKIMGLNSILW
PSWFISLILPLIVSAGLLVILKGNLLPYSDPSVVFVLSFAVAVTILQCFLSLTF
SRANLAACGGIITFYLYPVLVCAVQDYGFTLKFASLLSPVAFGFGCECFALFE
BOGIGVQNDNLFSVEEDGFNLTSVMMFLNTFLXGVTWYIEAVPGQYIGIPRW
YFPCYKSYFGESEDEKSHGPNQRISEICMEEPHHLKLGTVISQNLKVVYRDMKV
AVDGLANFYEGQITSLFHNGACKTTMTSLTGLTPTSTGTVILGKDIRSEMSTIR
QNLGVCPOHNVLFDMLTVEEHIWFYARLKGSKKVKAEOMEQALDGLPSSKLKST
SOLSGMOKLSVALAFVGGSKVLLDEPTAGVDPYXRRGIWELLKLRGRTILST
HMDPAVLDGRIALIIHSGKLCCVGGSEFLKQNLGTGYILVLYKKDVESLSSCRNS
STVYLKEDSVSSSSDAGLSDHSDTLTDIVSAISNLIRKHVSARLVEDIGHEL
TVILPYAAKEGAEVLFHEIDRDLGSSISYSETTLEETFLKVAESGVDAAETS
DGTLPARNRRAFQKQSLRPFTEDDAADPNDSDIDPESRETDLISGMDGKGSYQVK
GWKLTQQQFVALLKRLIIARRSRKFFAQILVPAVFCIALVPSLIVPPFGYPSLE
LOPMYNEQTFVSDAPEDTGTLELLNALTDPGCTRCMEGNPIPTPCOAGEEEM
TTAPVQOTIMDLFONGNWTQNPSPACQSSDKIKMLPVCPPGAGGLPPQKQNTA
DILQDLGRNISDLVITYVOIIRAKSLNKIWNFEFRYGGFSLGVSNTQALPPSEVN
DAIKOMKKHLKADSSADRLFNLGLRMTGLDRNNKVMFNKNGHAISSFLNIN
NAILRANLQGENPSHYGITAFNHLNLTQOLSEVALMTSDVLVSLVITFICFOQ
KSYVSTNLPLVALLLLYGWSITPLMYPASFVKIPSTAVVLTSLVNLFIGSVA
TFVLELTDNKLNINDILKSVLIFPHFCGLGRGLIDMKVQAMADALEREGERFVS
PLSWDLVGRNLFANAVEGVVFFLITVLIQIRFFIRPRVNAKISPLNDEDEVRREQ
RILGGQNDILETKELTKIYRRKRPFAVDRICVGPPIPGFGLLVNGAGKSTFKM
LTGDTVTYRGDAFLNKNLSINHEVHQNNGYCPQFDATITELTGRHEVFFALLRGV
PEKEVGKGMWAIKRLGVYGEYAGNYSNGNKKRLSTAMALITGGPPVFLDEPTTG
MDKARFLNLCALSVVKEGSRVYLTSHSMECEALCTRNAIMVNGFRCLGVSQHLK
NRFGDGTIVVRIAGSNPLDKPVQDFGLAFPGSVLKEKERNMLQYQLPSSLSLARI
FSILSQSKRLHLIEDYSVSQTLTDQVFNFAKQSDDDHLKDLISLHKNQTVADVAVLT
72893..72986
/gene="ABCA1"
/number=3
77520..77661
/gene="ABCA1"
/number=4
78931..79049
/gene="ABCA1"
/number=5
101048..101169
/gene="ABCA1"
/number=6
104152..104328
/gene="ABCA1"
/number=7
117287..117379
/gene="ABCA1"
/number=8
12329..122569
/gene="ABCA1"
/number=9
125285..125424
/gene="ABCA1"
/number=10
125754..125870
/gene="ABCA1"
/number=11
130079..130276
/gene="ABCA1"
/number=12
131023..131228
/gene="ABCA1"
/number=13
131749..131925
/gene="ABCA1"
/number=14
133708..133930

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version Replaced gi:6705871.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N_10

* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 872 971: contig of 871 bp in length
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length
* 3746 3845: gap of 100 bp
* 3846 4696: contig of 851 bp in length
* 4697 4796: gap of 100 bp
* 4797 5640: contig of 844 bp in length
* 5641 5740: gap of 100 bp
* 5741 6540: contig of 800 bp in length
* 6541 6640: gap of 100 bp
* 6641 7509: contig of 869 bp in length
* 7510 7609: gap of 100 bp
* 7610 8479: contig of 870 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 880 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length

* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length
* 20922 21021: gap of 100 bp
* 21022 21865: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 26621: contig of 851 bp in length
* 26622 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 27676: gap of 100 bp
* 27677 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp

Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Oshima, A.
 NEDO human cDNA sequencing project
 Unpublished (2000)
 2 (bases 1 to 1750)
 Isoqai, T. and Otsuki, T.
 Direct Submission
 Submitted (23-AGC-2000) to the DDBJ/EMBL/GenBank databases. Takao
 Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp,
 Tel: 81-438-52-3951, Fax: 81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing and clone selection:
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.

FEATURES
 Location/Qualifiers
 1..1750
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMA1000851"
 /clone_lib="MAMMA1"
 /tissue_type="Mammary gland"
 /notes="Cloning vector: pME18SFL3"
 BASE COUNT 291 a 489 c 586 g 384 t
 ORIGIN

Query Match 99.1% Score 219; DB 9; Length 1750;
 Best Local Similarity 100.0%; Pred. No. 2.1e-48;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 aattcgagcgagtgagtgaggccggagccgagcgagcgagccgtctctccc 62
 Db 1 AATTCGAGCGAGAGTGAGTGGGCGGGAGCCCGCAGAGCGGAGCCCTTCTCCC 60

Qy 63 gggctcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 122
 Db 61 GGGCTCGGCGAGCGAGCGGCGGGAGTCCGCGCACACAGAGCGGTTCTCAGGCGCG 120

Qy 123 ttgtctctgtttttcccccgggtctgtttctccctctctcggagcggtgtgtaagg 182
 Db 121 TTGCTCTGTTTTCCTCCCGGTTCTGTTTCTCCCTCTCCGAGGCGTTGTCAAG 180

Qy 183 ggtaggagaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 221
 Db 181 GGTAGGAGAAGAGAGCGCAACAAAGTGAAGAACAG 219

RESULT 15
 AC021345/C
 LOCUS Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
 AC021345
 ACCESSION AC021345
 VERSION AC021345.2 GI:9130845
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens, clone RP11-24J9
 Unpublished
 2 (bases 1 to 90698)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
 Boquslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,
 DeArellano, K., Dekar, K., Domino, M., Doyle, M., Fenestor, J.,
 Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
 Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., VO, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6705761.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submission@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4483
 Center clone name: 24_J_9

* NOTE: This record contains 92 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 910: contig of 910 bp in length
 * 911 1010: gap of 100 bp
 * 1011 1873: contig of 863 bp in length
 * 1874 1973: gap of 100 bp
 * 1974 2824: contig of 851 bp in length
 * 2825 2924: gap of 100 bp
 * 2925 3802: contig of 878 bp in length
 * 3803 3902: gap of 100 bp
 * 3903 4816: contig of 914 bp in length
 * 4817 4916: gap of 100 bp
 * 4917 5759: contig of 843 bp in length
 * 5760 5859: gap of 100 bp
 * 5860 6764: contig of 905 bp in length
 * 6765 6864: gap of 100 bp
 * 6865 7747: contig of 883 bp in length
 * 7748 7847: gap of 100 bp
 * 7848 8755: contig of 908 bp in length
 * 8756 8855: gap of 100 bp
 * 8856 9753: contig of 898 bp in length
 * 9754 9853: gap of 100 bp
 * 9854 10757: contig of 904 bp in length
 * 10758 10857: gap of 100 bp
 * 10858 11732: contig of 875 bp in length
 * 11733 11832: gap of 100 bp
 * 11833 12733: contig of 907 bp in length
 * 12740 12839: gap of 100 bp
 * 12840 13710: contig of 871 bp in length
 * 13711 13810: gap of 100 bp
 * 13811 14684: contig of 874 bp in length
 * 14685 14784: gap of 100 bp
 * 14785 15662: contig of 878 bp in length
 * 15663 15762: gap of 100 bp
 * 15763 16677: contig of 915 bp in length
 * 16678 16777: gap of 100 bp
 * 16778 17678: contig of 901 bp in length
 * 17679 17778: gap of 100 bp

TITLE
 JOURNAL
 COMMENT

QY 181 ggggtaggagaaagagacgcaaacacaaagtgaataacag 221
|||||
Db 188 ggggtaggagaaagagacgcaaacacaaagtgaataacag 228
|||||

RESULT 3
AAF92831
ID AAF92831 standard; DNA; 183999 BP.

AC AAF92831;

DT 17-MAY-2001 (first entry)

DE Human ABC1 genomic DNA.

XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.

KW Homo sapiens.

OS WO200115676-A2.

PN 08-MAR-2001.

PD 01-SEP-2000; 2000WO-IB01492.

PF 01-SEP-1999; 99US-0151977.

PR 15-MAR-2000; 2000US-0526193.

PR 23-JUN-2000; 2000US-0213958.

XX (UYB-) UNIV BRITISH COLUMBIA.

PA (XENO-) XENON GENETICS INC.

PI Hayden MR, Brooks-Wilson AR, Pimstone SM, Clee SM;

XX WPI; 2001-244356/25.

DR Treating a lower than normal high density lipoprotein-cholesterol
(HDL-C) level, a higher than normal triglyceride level, or a

PT cardiovascular disease, by administering a compound that modulates LXR-

PT or RXR-mediated transcriptional activity -

XX Claim 8; Fig 1; 317pp; English.

PS The present invention relates to a method for treating a patient

CC diagnosed as having a lower than normal high density

CC lipoprotein-cholesterol (HDL-C) level, a higher than normal

CC triglyceride level, or a cardiovascular disease, involving

CC administering a compound that modulates LXR- or RXR-mediated

CC transcriptional activity or ABC1 expression or activity.

CC The LXR gene product may be used in an assay to identify

CC compounds useful for the treatment of a disease or condition selected a

CC lower than normal HDL cholesterol level, a higher than normal

CC triglyceride level, and a cardiovascular disease.

XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

SQ
Query Match 99.8%; Score 220.6; DB 22; Length 183999;
Best Local Similarity 99.5%; Pred. No. 3.1e-54;
Matches 220; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 gtaattcgagcagagtagtgaggccgggaccccgagagccgagccgaccttcttc 60
|||||
Db 28790 gtaattcgagcagagtagtgaggccgggaccccgagagccgaccttcttc 28849
|||||
QY 61 ccgggctgcggcagggcagggcgggaggtcccgccaccaacagagccggttctcagggc 120
|||||
Db 28850 ccgggctgcggcagggcagggcgggaggtcccgccaccaacagagccggttctcagggc 28909
|||||
QY 121 gctttgctcttctttttcccggttctgttttctcccttctccggaagggttctcaa 180
|||||
Db 28910 gctttgctcttctttttcccggttctgttttctcccttctccggaagggttctcaa 28969
|||||

QY 181 ggggtaggagaaagagacgcaaacacaaagtgaataacag 221
|||||
Db 28970 ggggtaggagaaagagacgcaaacacaaagtgaataacag 29010
|||||

RESULT 4

AAH07432

ID AAH07432 standard; cDNA; 736 BP.

XX AAH07432;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:4267.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

KW Homo sapiens.

OS EP1074617-A2.

PN 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PF 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

PS Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13632 to AAH18742 represent human cDNA sequences; AAH92446 to

CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

SQ Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 99.3%; Score 219.4; DB 22; Length 736;

in gene therapy. The primers are useful for synthesising polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

XX Sequence 1556 BP: 380 A: 363 C: 399 G: 414 T: 0 other:
SO

[illegible][illegible]

XX 28 JUL 2000; 2000JUL 01182261.
XX
XX 29-JUL-1999; 99JP-0248036.
PR PR
XX 27-AUG-1999; 99JP-0300253.
PR PR
XX 11-JAN-2000; 2000JP-0118776.
PR PR
XX 02-MAY-2000; 2000JP-0183767.
PR PR
XX 09-JUN-2000; 2000JP-0241899.
PR PR
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI PI
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
PT
XX
XX Claim 1; SEQ ID 1564; 2537pp + CD ROM; English.
PS

The present invention describes primer sets for synthesising 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesising polynucleotides
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 AAH13633 to AAH18742 represent human cDNA sequences; AAB24446 to
 AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.

Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

Query Match	99.1%	Score	219;	DB	22;	Length	763;
Best Local Similarity	100.0%;	Pred. No.	1.7e-54;				
Matches	219;	Conservative	0;	Mismatches	0;	Indels	0;
Qy	3	aattcgagcagagtgagtggggccgagaccgcagagccgagccgagccctctctccc	62				
Db	1	aattcgagcagagtgagtggggccgagaccgcagagccgagccgagccctctctccc	60				
Qy	63	gggctgcgcagggcagggcgggagctccgcgcacacagagccggtttctcagggc	122				
Db	61	gggctgcgcagggcagggcgggagctccgcgcacacagagccggtttctcagggc	120				
Qy	123	tttgctcctgtttttcccggttctgttttccctcttccgaaagcgttgcaag	182				
Db	121	tttgctcctgtttttcccggttctgttttccctcttccgaaagcgttgcaag	180				
Qy	183	ggtaggagaaagagacgaacacaaaagtggaaaacag	221				
Db	181	qgtaggagaaagagacgaacacaaaagtggaaaacag	219				

RESULT	7	
AAH17451		
ID	AAH17451	standard; cDNA; 1750 BP.
XX	AC	
XX	AAH17451;	
XX		
DT	26-JUN-2001	(first entry)
XX		
XX	Human	cDNA sequence SEQ ID NO:16905.
DE		
XX	Human;	primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW		
XX	Homo sapiens.	
OS		
XX	EP1074617-A2.	
PN		
XX		
XX	07-FEB-2001.	
PD		
XX		
XX	28-JUL-2000;	2000EP-0116126.
PF		
XX		
XX	29-JUL-1999;	99JP-0248036.
PR		
XX	27-AUG-1999;	99JP-0300253.
PR		
PR	11-JAN-2000;	2000JP-0118776.
XX		
PR	02-MAY-2000;	2000JP-0183767.
PR		
PR	09-JUN-2000;	2000JP-0241899.
PR		

(HELI-) HELIX RES INST.

Ota T, Isodai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -

Claim 8; SEQ ID 16905; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.

Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;

	Query Match	99.1%	Score 219;	DB 22;	Length 1750;
	Best Local Similarity	100.0%;	Prod. No. 2.2e-54;		
	Matches 219;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3	aattgcagcagagtagtgaggccggagaccgcagagccgagccgaccttctctccc	62		
Db	1	aattgcagcagagtagtgaggccggagaccgcagagccgagccgaccttctctccc	60		
Qy	63	gggctgcgcgagcgaggccgggagctccgcgaccacagagccggttctcaggggcgc	122		
Db	61	gggctgcgcgagcgaggccgggagctccgcgaccacagagccggttctcaggggcgc	120		
Qy	123	tttgctcctgtttttcccggttctgttttctcccttctccgagaagcttgcgaagg	182		
Db	121	tttgctcctgtttttcccggttctgttttctcccttctccgagaagcttgcgaagg	180		
Qy	183	ggtaggagaagagacgcaaacacaaaagtggaaaacag	221		
Db	181	oatgagaagaagaacgcacaacacaaaagtgcgaataacag	219		

RESULT	8
AAK51683	
ID	AAK51683 standard; cDNA; 7281 BP.
XX	
XX	
AC	AAK51683;
XX	
XX	
DT	06-NOV-2001 (first entry)
XX	
XX	
DE	Human polynucleotide SEQ ID NO 228.
XX	
XX	
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW

OS Homo sapiens.

XX Key Location/Qualifiers

XX DT 291..7076

XX FT /*tag= a

XX FT /product= "ABCl polypeptide"

XX

XX WO200078972-A2.

XX PD 28-DEC-2000.

XX PF 16-JUN-2000; 2000WO-US16765.

XX PR 18-JUN-1999; 99US-0140264.

XX PR 14-SEP-1999; 99US-0153872.

XX PR 19-NOV-1999; 99US-0166573.

XX PA (CVTH-) CV THERAPEUTICS INC.

XX PI Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/14.

XX

XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide, useful for the development of agents for the treatment of heart disease and other disorders associated with hypercholesterolemia and atherosclerosis -

XX Disclosure; Page 122-128; 215pp; English.

XX The present sequence encodes a human adenosine triphosphate (ATP) binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell membranes and utilises ATP hydrolysis to transport a wide variety of substrates across the plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated mobilisation of intracellular cholesterol stores. ABC1 is defective in Tangier disease, a genetic disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins are useful for developing pharmaceutical agents for the treatment of heart disease and other disorders associated with hypercholesterolemia and atherosclerosis. The genes are useful for developing screening assays to screen for compounds that regulate the expression of genes associated with cholesterol transport. The genes and proteins are also useful for other disorders associated with hypercholesterolemia.

XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 89.1%; Score 197; DB 22; Length 10442;

Best Local Similarity 100.0%; Pred. No. 1e-47;

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggaccgcagagccgagccgaccttctctccgggctgcgagggcagggcg 84

Db 1 ggcgggaccgcagagccgagccgaccttctctccgggctgcgagggcagggcg 60

QY 85 ggagctccgcgacacagagccggtctctcagggcgcttgcctctttttcccg 144

Db 61 ggagctccgcgacacagagccggtctctcagggcgcttgcctctttttcccg 120

QY 145 gtctcttttctcccttctccgagcgttctcagggcgttagagaaagagacgcaaac 204

Db 121 gtctgtttctctcccttctccgagcgttctcagggcgttagagaaagagacgcaaac 180

QY 205 acaaaagtgaacacag 221

Db 181 acaaaagtgaacacag 197

RESULT 14

AAF24702

ID AAF24702 standard; DNA; 10442 BP.

XX AAF24702;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of a human ABC1 polypeptide.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1; apolipoprotein-mediated mobilisation; cholesterol; Tangier disease; chromosome 9q22-9q31; heart disease; hypercholesterolemia; atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 291..7076

XX FT /*tag= a

XX FT /product= "ABCl polypeptide"

XX WO200078971-A2.

XX PD 28-DEC-2000.

XX PF 16-JUN-2000; 2000WO-US16591.

XX PR 18-JUN-1999; 99US-0140264.

XX PR 14-SEP-1999; 99US-0153872.

XX PR 19-NOV-1999; 99US-0166573.

XX PA (CVTH-) CV THERAPEUTICS INC.

XX PA (UNIW) UNIV WASHINGTON.

XX Lawn RM, Wade D, Oram JF, Garvin M;

XX WPI; 2001-137811/14.

XX P-PSDB; AAB31365.

XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1 polynucleotides and polypeptides, useful for treatment of heart disease and other disorders associated with hypercholesterolemia and atherosclerosis -

XX Claim 3; Page 117-123; 211pp; English.

XX The present sequence encodes a human adenosine triphosphate (ATP) binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell membranes and utilises ATP hydrolysis to transport a wide variety of substrates across the plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated mobilisation of intracellular cholesterol stores. ABC1 is defective in Tangier disease, a genetic disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins are useful for developing pharmaceutical agents for the treatment of heart disease and other disorders associated with hypercholesterolemia and atherosclerosis. The genes are useful for developing screening assays to screen for compounds that regulate the expression of genes associated with cholesterol transport. The genes and proteins are also useful for other disorders associated with hypercholesterolemia.

XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 89.1%; Score 197; DB 22; Length 10442;

Best Local Similarity 100.0%; Pred. No. 1e-47;

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggaccgcagagccgagccgaccttctctccgggctgcgagggcagggcg 84

Db 1 ggcgggaccgcagagccgagccgaccttctctccgggctgcgagggcagggcg 60

QY 85 ggagctccgcgacacagagccggttctcagggcgcttgcctctttttcccg 144

Db 181 acaaaagtgaacacag 221

RESULT 14

AAF24702

ID AAF24702 standard; DNA; 10442 BP.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:21:31 ; Search time 110.16 Seconds
(without alignments)
492.783 Million cell updates/sec

Title: US-09-846-456-4

Perfect score: 221

Sequence: 1 gtaattgcagcgagatga.....aacacaaagtggaaaacag 221

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA:*
- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30.8	13.9	4480	3	US-09-191-171-7
C 2	30.8	13.9	4480	4	US-09-385-707-7
C 3	30.6	13.8	3509	2	US-08-327-832-4
C 4	30.6	13.8	3509	2	US-08-828-584-4
C 5	30.4	13.8	2067	2	US-08-713-928B-8
C 6	30.4	13.8	2155	3	US-09-191-171-4
C 7	30.4	13.8	2155	4	US-09-385-707-4
C 8	29.6	13.4	2255	4	US-08-871-572B-3
C 9	29.6	13.4	4405	1	US-07-885-972A-3
C 10	29.6	13.4	4405	2	US-08-745-880-3
C 11	29.6	13.4	4405	3	US-08-480-382-3
C 12	29.6	13.4	68750	3	US-09-335-409-1
C 13	29.6	13.4	68750	4	US-09-568-102-1
C 14	29.6	13.4	68750	4	US-09-567-969-1
C 15	29.6	13.4	68750	4	US-09-568-480-1
C 16	29.6	13.4	68750	4	US-09-568-486-1
C 17	29.6	13.4	68750	4	US-09-568-472-1
C 18	29.6	13.4	71989	4	US-09-443-501A-2
C 19	29.4	13.3	2699	3	US-09-167-354-2
C 20	29.2	13.2	1809	1	US-08-455-001-1
C 21	29.2	13.2	1809	5	PCR-US95-11869-1
C 22	29	13.1	1299	3	US-08-793-035-3
C 23	29	13.1	2754	4	US-09-429-322-3
C 24	28.8	13.0	30001	1	US-08-125-468-1
C 25	28.8	13.0	30001	2	US-08-474-933-1
C 26	28.8	13.0	4403765	4	US-09-103-840A-2
C 27	28.8	13.0	4411529	4	US-09-103-840A-1

28	28.6	12.9	2180	2	US-08-755-559-2	Sequence 2, Appli
29	28.6	12.9	2180	3	US-09-210-474-2	Sequence 2, Appli
30	28.6	12.9	2180	4	US-09-539-774-2	Sequence 2, Appli
31	28.6	12.9	20084	4	US-08-943-731-5	Sequence 5, Appli
C 32	28.4	12.9	6529	4	US-08-789-329C-1	Sequence 1, Appli
C 33	28.4	12.9	11219	1	US-07-642-734C-1	Sequence 1, Appli
C 34	28.4	12.9	11219	3	US-08-439-009A-1	Sequence 1, Appli
C 35	28.2	12.8	848	4	US-09-443-501A-5	Sequence 5, Appli
C 36	28.2	12.8	1470	1	US-08-570-311-3	Sequence 3, Appli
C 37	28.2	12.8	1470	2	US-08-353-485-3	Sequence 3, Appli
C 38	28.2	12.8	1841	2	US-08-570-311-5	Sequence 5, Appli
C 39	28.2	12.8	1841	2	US-08-353-485-5	Sequence 5, Appli
C 40	28.2	12.8	3789	1	US-07-872-644-42	Sequence 5, Appli
C 41	28.2	12.8	3789	1	US-08-297-494-4	Sequence 42, Appl
C 42	28.2	12.8	3789	1	US-08-297-510-42	Sequence 42, Appl
C 43	28.2	12.8	3789	1	US-08-479-532-42	Sequence 42, Appl
C 44	28.2	12.8	3789	1	US-08-455-526-42	Sequence 42, Appl
C 45	28.2	12.8	3789	1	US-08-455-525-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-191-171-7/c
; Sequence 7, Application US/09191171
; Patent No. 6149909
; Patent No. 6149909 6143294
; GENERAL INFORMATION:
; APPLICANT: SCOTT, Hamish S.
; APPLICANT: ANSON, Donald S.
; APPLICANT: ORSBORN, Annette M.
; APPLICANT: NELSON, Paul V.
; APPLICANT: CLEMENTS, Peter R.
; APPLICANT: MORRIS, Charles P.
; APPLICANT: HOPWOOD, John J.
; TITLE OF INVENTION: SYNTHETIC ALPHA-L-IDURONIDASE AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,171
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/494,104
; FILING DATE: 23-JUN-1995
; APPLICATION NUMBER: AU PK9490/91
; FILING DATE: 14-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU92/00611
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,254
; FILING DATE: 07-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 89782
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366

US-08-327-832-4

APPLICANT: Wasley, Louis
APPLICANT: Wong, Polly A

APPLICANT: Kautman, Rnadal J.
APPLICANT: Tekamp-Olson, Patricia

; APPLICANT: Wong, Polly

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:14:51 ; Search time 5225.75 Seconds
(without alignments)
636.716 Million cell updates/sec

Title: US-09-846-456-5

Perfect score: 159

Sequence: 1 ttaatgaccagccacggcg.....cttcagaagaagacaaaca 159

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htggo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

1	159	100.0	159	6	AX351033	Sequence
2	159	100.0	357	6	AX351030	Sequence
3	155.8	98.0	10442	6	AX060713	Sequence
4	155.8	98.0	10442	6	AX060892	Sequence
5	155.8	98.0	10442	9	AF285167	Homo sapi
6	155.8	98.0	10474	6	AX060719	Sequence
7	155.8	98.0	10474	6	AX060721	Sequence
8	155.8	98.0	10474	6	AX060898	Sequence
9	155.8	98.0	10474	6	AX060900	Sequence
10	155.8	98.0	149034	9	AF275948	Homo sapi
11	153.2	96.4	183999	6	AX092589	Sequence
12	143.4	90.2	697	9	AF258627	Homo sapi
13	142.2	89.4	200	9	AF258623S2	Homo sapi
14	142.2	89.4	446	6	AX127764	Sequence
15	142.2	89.4	446	6	AX127764	Sequence
16	142.2	89.4	480	9	HS252277	Homo sapi
17	142.2	89.4	7260	6	AX253452	Sequence
18	142.2	89.4	9741	6	AX127830	Sequence
19	142.2	89.4	9741	6	AX139817	Sequence
20	142.2	89.4	9741	6	AX351038	Sequence
21	142.2	89.4	9854	6	AX127831	Sequence
22	142.2	89.4	9854	6	AX139818	Sequence
23	142.2	89.4	129608	9	AL353685	Human DNA
24	142.2	89.4	175064	2	AC012230	Homo sapi
25	142.2	89.4	201144	9	AF287262	Homo sapi
26	140.6	88.4	1556	9	AK024328	Homo sapi
27	136.2	85.7	298	9	AB037924	Homo sapi
28	135.2	85.0	7860	6	AB092594	Sequence
29	135.2	85.0	7862	6	AX135712	Sequence
30	130.2	81.9	90698	2	AC021345	Homo sapi
31	72.8	45.8	207659	2	AC091466	Mus muscu
32	72.8	45.8	278572	10	AF287263	Mus muscu
33	67	42.1	7878	10	MMABCL	Mus muscu
34	64.4	40.5	6786	9	AB055982	Homo sapi
35	54	34.0	7074	5	AF362377	Gallus ga
36	37	23.3	110000	2	LMFLCHR36_08	Continuation (9 of
37	35.8	22.5	61524	9	AL353725	Human DNA
38	35.6	22.4	52376	2	AC102919	Mus muscu
39	35.4	22.3	37	6	AX092843	Sequence
40	35	22.0	22881	10	AF287142	Mus muscu
41	35	22.0	23024	6	AX080494	Sequence
42	35	22.0	185825	10	AC087114	Mus muscu
43	35	22.0	200143	2	AC073805	Mus muscu
44	34.6	21.8	6704	9	AF328787	Homo sapi
45	34.6	21.8	6768	6	AX320362	Sequence

ALIGNMENTS

RESULT	1	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002	
DEFINITION	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002	
ACCESSION	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002	
VERSION	AX351033.1	GI:18616389						
KEYWORDS	human.							
SOURCE	ORGANISM	human.						
		Homo sapiens						
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;						
		Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.						
REFERENCE	1	(sites)						
AUTHORS		Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P.,						
		Brewer,B., Duverger,N., Renaley,A. and Santamarina-Fojo,S.						
TITLE		Regulatory nucleic acid sequences of the abcl gene						
JOURNAL		Patent: WO 0183746-A 5 08-NOV-2001;						
		Aventis Pharma S.A. (FR)						
FEATURES		Location/Qualifiers						
		1..159						
		/organism="Homo sapiens"						
		/db_xref="taxon:9606"						
BASE COUNT		32 a	47 c	48 g	32 t			
ORIGIN								

```

Query Match      100.0%; Score 159; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.7e-33;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttaatgaccagccagggcgctccctctctcagctctgcccgtcctccagggctccc 60
Db 1 TTAATGACCAGCCAGGGCGCTCCCTCTGCTGACGTCTGGCCGCTGCTTCCAGGGCTCCC 60

QY 61 gagccacacgctggcgctgctgaggaacatggcatgttgctcagctgaggttg 120
Db 61 GAGCCACACGCTGGCGTGTGCTGAGGAACATGGCATGTGGCTCAGCTGAGGTG 120

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
Db 121 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 159

RESULT 2
AX351030 AX351030 357 bp DNA linear PAT 06-FEB-2002
LOCUS AX351030 Sequence 2 from Patent WO0183746.
DEFINITION AX351030
ACCESSION AX351030
VERSION AX351030.1 GI:18616386
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Rosier-Montus M.F., Prades C., Lemoine C., Naudin L., Denefle P.,
Brewer B., Duverger N., Remaley A. and Santamarina-Fojo S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 2 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
1. .357
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 64 a 78 c 116 g 99 t
ORIGIN

Query Match      100.0%; Score 159; DB 6; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.6e-33;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttaatgaccagccagggcgctccctctctcagctctgcccgtcctccagggctccc 60
Db 100 TTAATGACCAGCCAGGGCGCTCCCTCTGCTGACGTCTGGCCGCTGCTTCCAGGGCTCCC 159

QY 61 gagccacacgctggcgctgctgaggaacatggcatgttgctcagctgaggttg 120
Db 160 GAGCCACACGCTGGCGTGTGCTGAGGAACATGGCATGTGGCTCAGCTGAGGTG 219

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
Db 220 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 258

RESULT 3
AX060713 AX060713 10442 bp DNA linear PAT 22-JAN-2001
LOCUS AX060713 Sequence 1 from Patent WO0078972.
DEFINITION AX060713
ACCESSION AX060713
VERSION AX060713.1 GI:12406103
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE Atp binding cassette transporter protein abcl1 polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
1. .10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t
ORIGIN

Query Match      98.0%; Score 155.8; DB 6; Length 10442;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttaatgaccagccagggcgctccctctctcagctctgcccgtcctccagggctccc 60
Db 198 TTAATGACCAGCCAGGGCGCTCCCTCTGCTGAGCTCTGGCCGCTGCTTCCAGGGCTCCC 257

QY 61 gagccacacgctggcgctgctgaggaacatggcatgttgctcagctgaggttg 120
Db 258 GAGCCACACGCTGGCGTGTGCTGAGGAACATGGCTTGTGGCTCAGCTGAGGTG 317

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
Db 318 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 356

RESULT 4
AX060892 AX060892 10442 bp DNA linear PAT 22-JAN-2001
LOCUS AX060892 Sequence 1 from Patent WO0078971.
DEFINITION AX060892
ACCESSION AX060892
VERSION AX060892.1 GI:12406270
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE Atp binding cassette transporter protein abcl1 polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
1. .10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t
ORIGIN

Query Match      98.0%; Score 155.8; DB 6; Length 10442;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttaatgaccagccagggcgctccctctctcagctctgcccgtcctccagggctccc 60
Db 198 TTAATGACCAGCCAGGGCGCTCCCTCTGCTGAGCTCTGGCCGCTGCTTCCAGGGCTCCC 257

QY 61 gagccacacgctggcgctgctgaggaacatggcatgttgctcagctgaggttg 120
Db 258 GAGCCACACGCTGGCGTGTGCTGAGGAACATGGCTTGTGGCTCAGCTGAGGTG 317

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
Db 318 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 356

RESULT 5
AF285167 AF285167 10442 bp mRNA linear PRI 09-AUG-2000
LOCUS AF285167

```

DEFINITION Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA, complete cds.
 ACCESSION AF285167
 VERSION AF285167.1 GI:9755158
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 10442)
 AUTHORS Schwartz, K., Lawn, R.M. and Wade, D.P.
 TITLE ABCA1 gene expression and apolipoprotein A-I-mediated cholesterol efflux are regulated by LXR
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 10442)
 AUTHORS Lawn, R.M., Wade, D.P., Garvin, M.R., Wang, X., Schwartz, K., Porter, J.G., Seilhamer, J.J., Vaughan, A.M. and Oram, J.F.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUL-2000) Discovery Research, CV Therapeutics Inc., 3172 Porter Drive, Palo Alto, CA 94304, USA
 FEATURES
 1. Location/Qualifiers
 1..10442
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q31"
 /cell_type="fibroblast"
 /tissue_type="skin"
 1..10442
 /gene="ABCA1"
 291..7076
 /gene="ABCA1"
 /codon_start=1
 /product="ATP-binding cassette transporter 1"
 /protein_id="AAF98175.1"
 /db_xref="GI:9755159"
 /translation="MACPQRLLLKRLNLTFRRTQQLLEVAWPLFIPLILSVRL
 SYPEGEHCHFPNKAQPSAGTLPWOGIICNANNPCFRPTCEAPGVGNFNKSI
 ARFLSDARRLLYSOKDTSMKMRKVLRTQIKKSSNLKLODFLVDNTEFSGLVH
 NLSLPKSTVDKMLRADVLHKVFLQGLYGLHSLCNGSKSEEMQLQDQVSELCGLP
 KEKLAARVLRSLNDILKFLRLNSTSPFKELAEATKILSLGTLGAELEFNSR
 SWSMDREVFMTNWNSSSTQIYQAVSRIVCGHGGGKIKKSLNMYEDNNYKALF
 GNGTEDEATFYDNTTPYCNLNMKNLSPSLRIIWKALKPLVGLIKILYTPDTPAT
 RQVMAEYNTKTELAVFLHLEGGMEELSPKTIWFMENSOEMLVRLMDRDNHFEW
 QQLDGLTWAQDIIVAFKAPEDVOSSGVIYTWREAFNETNOAIRTISREMECVNLN
 KLEPIATEVNLINKSMELLDERKFMAGIVGTITPGSIELPHVKKYKIRMDINVERT
 NKIKDGYWDGPRADPRDMRYVWGFAYLQDVVEQALIRVLGTEKTKGVYMQMPY
 PCYVDDIFLRVMSMPLFMFLAWIYSVAVILIRGIVYEKEARLKTMRINGLNDSTILW
 FSWFISLLIPLLVAGLLVILKGLNLLPYSDPSVVFVLSVFAVVTILQCFILSTILF
 SRANLAAACGGIYFTLLYPVLCVAMQDVGFTLKFASLLSPVAFPGCYGFLPFE
 EOGIGVQWMDNLFSPVEDGFNLTSTISMLFDTFLYGVMTWVIEAVFPGCYGIPRPW
 YFCTCKSYRFEESDEKSHPSNOKRMESECEEPHKLGVSIONLVKVRDGMKV
 AVDGLALNFYEGQITSLFHNGAGKTTMTSILGLFPPTSGTAYILGKDIRSEMTIR
 QNLGVCQHNVLFDMLVEGHIFYARKLGLSEKHKVKAEMEQLDGLPSSKLKSTK
 SQSGMQGKLSVALAEVGSKVILDEPTAGVDPSYRGGIWEILLKLYRGRTIILST
 HMDAEADVLDRITAIISGKLCVGSLLFNQGLTGYTLTLVKDVESSLCRNS
 STVYLKKEQSDVSQSDAGLSDHSDTLTIDVSAISNLIRKHSRPARVEDIGHLEL
 TYVLPYEAAGAGAVFELFHEIDRLSDGLSSYGISSETLLEELFLKVAESGYDAETS
 DGLTPARNRRAFGDKQSLRPETEDDAADPNSDIDPESRETDLLSGMDGKGSYQVK
 GWKLTQQFVALLMKRLLIARRSKGFFAIVLPVAFVCIALVFLSIVPFGYKPSLE
 LQPMYNEQYTFVNSNADPDTGTLELLNALTDPFGFTRCMEGNIPDTPCQAGESEW
 TTAPOQTIMDLFONGWNTWQNSPACQSSDKIKMLPVPVPGAGGLPPQKQNTA
 DILQDLTGRNIDSLVLYVYQIIAKSLKNKIWNFEYRGGFSLGVSNTQALPSPQVFN
 DAIRKQMKHLKLAKDSADREFLNSLGFMTGLDTRNNVKNFNNKNGHAISSFLNVLN
 NAILRANLQGENPSHGKIFAFNPLNLTKQLESEVALMTTISVDLVLSVICIFAMSFV
 PASVFEILQERSVKARHLOFISGVKPIVYLSNFMVDMCNVYVPATLVIIIFICFQ
 KSYVSSNTLPLVALLLLLYLGNSTPLMPASFFVKITPSTAYVLTYSNLTIGNSVFA
 TFLVLELTDNKNLINDILKSVLIFPFCGLRGILDMKNQAMADALERGENRVS
 RLSWDLVRNLFAMAVGVVFFLITVLIQYRFRPRPVNAKLSPLNDEDEDVRRRQ
 PLDGGQNDILELKELTLYRKRKPAVDRIKCVIPPGECFGLLVNGAGKSTFKM
 LTGDTTVTRGDALFNKNSILSNIEHVHQNMGICPQFDATTELLTTRGVHFFALLRGV
 PERKVGKGMWAIKGLVYKGYAGNYSGGNKKRLSTAMALIGPPVYFLDPEPTG
 MDPKARRELMNCAISVYKGRSVVLTSHSMECEALCTRNAIMVNGFRCLGSVOHLK

gene

CDS

BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
 ORIGIN
 Query Match 98.0%; Score 155.8; DB 9; Length 10442;
 Best Local Similarity 98.7%; Pred. No. 1.le-32;
 Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 ttaatgaccagccagggcgctgcctgctgcagctctgagcctgagcctccagggctccc 60
 |||||
 Db 198 TTAATGACCAGCCAGCGCGCTCCTGCTGTGAGCTCTGCGCGCTGCCTTCAGGGCTCCC 257
 |||||
 Qy 61 gagccacacgctggcgctgctgctgagggagaaatgagcatgttggtgcctagctgaggttg 120
 |||||
 Db 258 GAGCCACACGCTGGCGCTGCTGCTGAGGGAACATGCTTGTGGCCTCAGCTGAGGTG 317
 |||||
 Qy 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 Db 318 CTGCTGTGGAGAACCTTCTACCTTCAGAGAGACAAACA 356
 |||||
 RESULT 6
 AX060719
 LOCUS AX060719 10474 bp DNA linear PAT 22-JAN-2001
 DEFINITION Sequence 7 from Patent WO0078972.
 ACCESSION AX060719
 VERSION AX060719.1 GI:12406108
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 10474)
 AUTHORS Lawn, R.M., Wade, D. and Garvin, M.
 TITLE Regulation with binding cassette transporter protein abc1
 JOURNAL Patent: WO 0078972-A 7 28-DEC-2000;
 CV THERAPEUTICS, INC. (US)
 FEATURES
 Location/Qualifiers
 1..10474
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others
 ORIGIN
 Query Match 98.0%; Score 155.8; DB 6; Length 10474;
 Best Local Similarity 98.7%; Pred. No. 1.le-32;
 Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 ttaatgaccagccagggcgctgcctgctgcagctctgagcctgagcctccagggctccc 60
 |||||
 Db 230 TTAATGACCAGCCAGCGCGCTCCTGCTGTGAGCTCTGCGCGCTGCCTTCAGGGCTCCC 289
 |||||
 Qy 61 gagccacacgctggcgctgctgctgagggagaaatgagcatgttggtgcctagctgaggttg 120
 |||||
 Db 290 GAGCCACACGCTGGCGCTGCTGCTGAGGGAACATGCTTGTGGCCTCAGCTGAGGTG 349
 |||||
 Qy 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 Db 350 CTGCTGTGGAGAACCTTCTACCTTCAGAGAGACAAACA 388
 |||||
 RESULT 7
 AX060721
 LOCUS AX060721 10474 bp DNA linear PAT 23-JAN-2001
 DEFINITION Sequence 9 from Patent WO0078972.
 ACCESSION AX060721
 VERSION AX060721.1 GI:12406109
 KEYWORDS human.
 SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 10474)
TITLE Regulation with binding cassette transporter protein abcl
JOURNAL Patent: WO 0078972-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
Source 1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN
Query Match 98.0%; Score 155.8; DB 6; Length 10474;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ttaatgaccagcagcggtccctgctgctgagcctggtgagcctgctccctccagggtccc 60
|||||
Db 230 TTAATGACCAGCAGCGGGCTCCCTGCTGCTGAGCTGTGGCGCTGCTTCCAGGCTCCC 289
Qy 61 gagccacacgctggcgctgctgctgctgaggaacatggtgctgctcagctgaggttg 120
|||||
Db 290 GAGCCACACGCTGGCGTGGCTGGCTGAGGAACATGGTGTGTGGCCTCAGCTGAGGTG 349
Qy 121 ctgctgtggaagaacctcatttcagaagaagacaaca 159
|||||
Db 350 CTGCTGTGGAAGAACCCTCACTTTCAGAAGAAGACAACA 388
RESULT 8
AX060898
LOCUS AX060898 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 7 from Patent WO0078971.
ACCESSION AX060898
VERSION AX060898.1 GI:12406275
KEYWORDS human.
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 10474)
TITLE Regulation with binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 7 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
Source 1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others
ORIGIN
Query Match 98.0%; Score 155.8; DB 6; Length 10474;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ttaatgaccagcagcggtccctgctgctgagcctggtgagcctgctccctccagggtccc 60
|||||
Db 230 TTAATGACCAGCAGCGGGCTCCCTGCTGAGCTGTGGCGCTGCTTCCAGGCTCCC 289
Qy 61 gagccacacgctggcgctgctgctgctgaggaacatggtgctgctcagctgaggttg 120
|||||
Db 290 GAGCCACACGCTGGCGTGGCTGGCTGAGGAACATGGTGTGTGGCCTCAGCTGAGGTG 349
Qy 121 ctgctgtggaagaacctcatttcagaagaagacaaca 159
|||||
Db 350 CTGCTGTGGAAGAACCCTCACTTTCAGAAGAAGACAACA 388
RESULT 9
AX060900
LOCUS AX060900 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078971.
ACCESSION AX060900
VERSION AX060900.1 GI:12406276
KEYWORDS human.
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 10474)
TITLE Atp binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
Source 1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN
Query Match 98.0%; Score 155.8; DB 6; Length 10474;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ttaatgaccagcagcggtccctgctgctgagcctggtgagcctgctccctccagggtccc 60
|||||
Db 230 TTAATGACCAGCAGCGGGCTCCCTGCTGCTGAGCTGTGGCGCTGCTTCCAGGCTCCC 289
Qy 61 gagccacacgctggcgctgctgctgctgaggaacatggtgctgctcagctgaggttg 120
|||||
Db 290 GAGCCACACGCTGGCGTGGCTGGCTGAGGAACATGGTGTGTGGCCTCAGCTGAGGTG 349
Qy 121 ctgctgtggaagaacctcatttcagaagaagacaaca 159
|||||
Db 350 CTGCTGTGGAAGAACCCTCACTTTCAGAAGAAGACAACA 388
RESULT 10
AF275948
LOCUS AF275948 149034 bp DNA linear PRI 17-JUL-2000
DEFINITION Homo sapiens ABCA1 (ABCA1) gene, complete cds.
ACCESSION AF275948
VERSION AF275948.1 GI:9247085
KEYWORDS human.
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 149034)
TITLE Santamarina-Fojo, S., Peterson, K., Knapper, C., Qiu, Y., Freeman, L.,
Cheng, J.F., Osorio, J., Remaley, A., Yang, X.P., Haudenschild, C.,
Prades, C., Chimini, G., Blackmon, E., Francois, T., Duverger, N.,
Rubin, E.M., Rosier, M., Deneffe, P., Fredrickson, D.S. and Brewer, H.B.
Jr.
Complete genomic sequence of the human ABCA1 gene: analysis of the
human and mouse ATP-binding cassette A promoter
Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
JOURNAL 20345099
MEDLINE
REFERENCE 2 (bases 1 to 149034)
AUTHORS Santamarina-Fojo, S., Peterson, K.M., Knapper, C.L., Freeman, L.A.,
Remaley, A.T., Yang, X.-P., Haudenschild, C.C., Blackmon, E.E.,
Francis, T.L. and Brewer, H.B. Jr.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA
FEATURES
Location/Qualifiers
Source 1..149034

mRNA	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	Join(1454, .1674, 25831, .25989, 40385, .40478, 45012, .45153, 46423, .46541, 67729, .67850, 70831, .71007, 83960, .84052, 89102, .89250, 91962, .92101, 92433, .92549, 96758, .96955, 97702, .97907, 98428, .98604, 100391, .100613, 102360, .102581, 103642, .103846, 104951, .105064, 106862, .107033, 108023, .108154, 109460, .109602, 109806, .109943, 110646, .110866, 112124, .112196, 113183, .113385, 115053, .115101, 115297, .115410, 116806, .116954, 118604, .118728, 119964, .120062, 123094, .123283, 124804, .124988, 126208, .126240, 127363, .127468, 128943, .129017, 129539, .129708, 130939, .131116, 133114, .133229, 133341, .133485, 134526, .134649, 135736, .135865, 136130, .136250, 137038, .137100, 138008, .138114, 140469, .140610, 140982, .141116, 142060, .142163, 142646, .142738, 143397, .143640, 144581, .148034)		
	/gene="ABCA1"		
	/product="ABCA1"		
	5' UTR	Join(1454, .1674, 25831, .25923)	
	gene	/gene="ABCA1"	1454, .148034
	repeat_region	/rpt_family="Alu"	2406, .2678
	repeat_region	/rpt_family="Alu"	4240, .4509
	repeat_region	/rpt_family="Alu"	4998, .5270
	repeat_region	/rpt_family="Alu"	5563, .5839
	repeat_region	/rpt_family="Alu"	6615, .6877
	repeat_region	/rpt_family="Alu"	8800, .9046
	repeat_region	/rpt_family="Alu"	9998, .10279
	repeat_region	/rpt_family="Alu"	11859, .12133
	repeat_region	/rpt_family="Alu"	12810, .12902
	repeat_region	/rpt_family="Alu"	15220, .15399
	repeat_region	/rpt_family="Alu"	15508, .15787
	repeat_region	/rpt_family="Alu"	18600, .18860
	repeat_region	/rpt_family="Alu"	20950, .21206
	repeat_region	/rpt_family="Alu"	25038, .25314
	CDS	/rpt_family="Alu"	Join(25924, .25989, 40385, .40478, 45012, .45153, 46423, .46541, 67729, .67850, 70831, .71007, 83960, .84052, 89102, .89250, 91962, .92101, 92433, .92549, 96758, .96955, 97702, .97907, 98428, .98604, 100391, .100613, 102360, .102581, 103642, .103846, 104951, .105064, 106862, .107033, 108023, .108154, 109460, .109602, 109806, .109943, 110646, .110866, 112124, .112196, 113183, .113385, 115053, .115101, 115297, .115410, 116806, .116954, 118604, .118728, 119964, .120062, 123094, .123283, 124804, .124898, 126208, .126240, 127363, .127468, 128943, .129017, 129539, .129708, 130939, .131116, 133114, .133229, 133341, .133485, 134526, .134649, 135736, .135865, 136130, .136250, 137038, .137100, 138008, .138114, 140469, .140610, 140982, .141116, 142060, .142163, 142646, .142738, 143397, .143640, 144581, .144721)
		/codon_start=1	
		/product="ABCA1"	
		/protein_id="AAF86276.1"	
		/db_xref="GI:9247086"	
		/translation="MACWPQLRLLNKNTFRRRTQCQLLLEVAWPLFLLISVRL	

Qy 1

Db 212 CTGCTGTGGAGAACCTCAGCTTTCAGAGAGACAAACA 250

Search completed: September 20, 2002, 03:15:13
Job time: 14027 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:14 ; Search time 3895 seconds
(without alignments)
765.809 Million cell updates/sec

Title: US-09-846-456-4
Perfect score: 221
Sequence: 1 gtaattgcgcgagagtgta.....aacacaaagtggaacacag 221

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_estl.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219.4	99.3	736	9	AU135588
2	219	99.1	763	9	AU121731
3	204	92.3	222	10	Z44377
4	178.2	80.6	998	10	BG678861
5	67.4	30.5	547	12	AZ769996
6	67.4	30.5	619	9	BB657864
7	65.4	29.6	625	9	BB665939
8	57.8	26.5	218	9	BB594197
9	57.8	26.2	259	9	BB594265
10	55.2	25.0	272	9	BB568993
11	50.4	22.8	276	9	BB570397
12	47.4	21.4	535	10	BG384217
13	39.2	17.7	925	12	CNS0091P
14	38.8	17.6	811	12	BH470300
15	38.4	17.4	910	12	CNS0060N
16	37.4	16.9	879	12	CNS020OG
17	37.2	16.8	311	9	BB374442

18	37.2	16.8	594	10	C76305
19	37.2	16.8	1051	12	AG127909
c 20	37	16.7	700	10	BF237183
c 21	37	16.7	981	12	AG131276
c 22	36.8	16.7	315	12	CNS01677
c 23	36.8	16.7	1015	12	AG137074
c 24	36.6	16.6	313	9	BB318923
c 25	36.4	16.5	904	10	BI101161
c 26	36.2	16.4	353	9	AI850822
c 27	36.2	16.4	429	9	AW489757
c 28	36.2	16.4	457	9	AW490154
c 29	36.2	16.4	466	10	BE948350
c 30	36.2	16.4	476	9	BB698363
c 31	36.2	16.4	489	10	BE943674
c 32	36.2	16.4	489	10	BE943841
c 33	36.2	16.4	493	10	BE947468
c 34	36.2	16.4	557	9	BB696895
35	36.2	16.4	650	9	AV334840
36	36.2	16.4	656	9	BB547934
c 37	36	16.3	1158	12	AG073892
c 38	35.6	16.1	320	9	BB317508
c 39	35.6	16.1	400	10	BB692826
c 40	35.6	16.1	409	10	BI082717
c 41	35.6	16.1	411	9	AA098273
c 42	35.6	16.1	411	9	BB790023
c 43	35.6	16.1	442	10	BF224565
c 44	35.6	16.1	442	10	BF319828
c 45	35.6	16.1	446	9	AA562016

ALIGNMENTS

RESULT 1
AU135588
LOCUS AU135588 PLACE1 Homo sapiens cDNA clone PLACE1002437 5', mRNA linear EST 24-OCT-2000
DEFINITION AU135588 736 bp
ACCESSION AU135588.1 GI:10996127
VERSION AU135588.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 736)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
source location/Qualifiers
1..736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"
/tissue_type="placenta"
/note="Vector: pME18SFL3"
163 a 199 c 199 g 170 t 5 others
BASE COUNT
ORIGIN

1051 12 AG127909
700 10 BF237183
981 12 AG131276
315 12 CNS01677
1015 12 AG137074
313 9 BB318923
904 10 BI101161
353 9 AI850822
429 9 AW489757
457 9 AW490154
466 10 BE948350
476 9 BB698363
489 10 BE943674
489 10 BE943841
493 10 BE947468
557 9 BB696895
650 9 AV334840
656 9 BB547934
1158 12 AG073892
320 9 BB317508
400 10 BB692826
409 10 BI082717
411 9 AA098273
411 9 BB790023
442 10 BF224565
442 10 BF319828
446 9 AA562016

Qy 77 cagggcgggagctccgcgcaccaacagagccggttctcagggcgctttgctccttgttt 136

Db	61	CAGGCGGGAGCTCCGCGACCAACAGACCGGTTCTCAGGGCGCCTTGCCTCTGT	120
Qy	137	tttcccggtctgtttttctcccocttccggaagcttgcgaagggtaggaagaag	196
Db	121	TTTCCCGGTTCTGTTTCTCCCTTTCGGAAGGCTGTCAAGGGTAGGAAAGAG	180
Qy	197	acgcaaacacaaagtggaaaaag	221
Db	181	ACGCAAAACACAAAGTGAAACAG	205

RESULT	4	
LOCUS	BG6778861	
DEFINITION	60624760F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735', mRNA sequence.	
ACCESSION	BG6778861	998 bp mRNA linear EST 01-MAY-2001
VERSION	BG678861.1	GI:113910258
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 998)	
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov Tissue procurement: James Cleaver, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10603 row: g column: 16 High quality sequence stop: 860.	

[illegible]

RESULT	5
AZ769996	
LOCUS	
DEFINITION	AZ769996 547 bp DNA linear GSS 16-FEB-2001 clone UGUCIM0571A17 F, DNA sequence.
ACCESSION	AZ769996
VERSION	AZ769996.1 GI:12890721
KEYWORDS	GSS.
SOURCE	house mouse
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 547)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly, M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished (2000)
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT, 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0571 row: A column: 17 Seq primer: CGTGTAAAACGACGGCCAGT Class: plasmid ends High quality sequence stop: 547.

[illegible]

FEATURES	SOURCE
----------	--------

BASE COUNT	120 a	61 c	61 g	172 t	511 others
------------	-------	------	------	-------	------------

Matches	32;	Conservative	86;	Mismatches	68;	Indels	0;	Gaps	0;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

Qy 66 ctgcqccagggcaggcgaggctccgacccaacagagccggttctcagggcgttt 125

Ov		126	gqtccttgrttttttccrcgggtctcgttttctccccctctcccccgagatctccccccc+	185
----	--	-----	---	-----

0v 186 aqaa 191

RESULT	14
BH470300	
LOCUS	BH470300 . 811 bp DNA linear GSS 13-DEC-2001
DEFINITION	BOHMIG9TR BOHM Brassica oleracea genomic clone BOHMIG9, DNA sequence.

ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica

TITLE	Whole genome shotgun sequencing of <i>Brassica oleracea</i>
JOURNAL	Unpublished (2001)
COMMENT	Other_GSSs: BOHMI69TF Contact: Chris Town

Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

```

FEATURES
source
Location/Qualifiers
1. .811
/organism="Brassica oleracea"

```

```

/clone_lib="BOHM"
/note="Vector: pHS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHS1 using BstXI linkers"

```

BASE COUNT	241 a	122 c	178 g	270 t	FROM LEADY SOURCE MARKERS
ORIGIN					

Ov 11 acgagagtgaagtggggccgggagcccgacagagccgagcccttctctccgggactaca 70

Qy 71 gcagggcaggcgggagctccgcgcacacagcgcgattctcagggcgctttgctcc 130
Dp 136 GTACGCGTTTGTGGCGCCAAATATGCCCACTGCTGCGACGACATGCTGCATTTATCAGC 195

Db 196 CTCGTTTTACCGCAGTTTTCTGACAACGCGTTTCTGCTAGGATGGTCAAGGTGAAGAGG 255

Ov 191 aaagagacacaaac 204

RESULT 15
CNS006ON
LOCUS

	DNA	LOCUS	CSS 02-TUN-1999
CNS006ON	910 bp		

fly), genomic survey sequence.
AL065629
ACCESSION
VERSION
AL065629.1 GI:4944698
KEYWORDS
GSS

REFERENCE
1. Gerasimov, I. G. 1910. *Tracheata, Insecta; Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephýdroidea; Drosophilidae; Drosophila*. 1, 683 pp. 10, 010.

JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
Web : www.genoscope.cns.fr
National Institute of Biological Sciences
COMMENT

melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Kazuo Yamawaki, Drosophila Genome Project, The Institute of

ECORI digestion of *Drosophila* DNA provided by the BDCP from the isogenic strain y2; cn bw sp, the same strain used for the BDCP's Pl and EST libraries. A more detailed description of the library and how to order it is available at <http://www.bdc.org>.

FEATURES	source
Location/Qualifiers	
1. .910	
/organism="Drosophila melanogaster"	
/start	

BASE COUNT	202 a	63 C	112 g	198 t	335 others
ORIGIN	/note="end : T7"				

Best Local Similarity	18.9%	pred. NO. 28;	
Matches	34;	Conservative	30;
		Mismatches	66;
		Indels	0;
		Gaps	0;

9 gacgagagtgagtggggccggaccgcagagccgagccgaccttctctcccggtg 68

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:09:16 ; Search time 520.94 Seconds
(without alignments)
524.032 Million cell updates/sec

Title: US-09-846-456-5

Perfect score: 159

Sequence: 1 ttaatgaccacgccacggcg.....ctttcagaagaagacaaca 159

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL1/gcgdata/hold-geneseg/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155.8	98.0	10442	22	AAF24680 Nucleotide sequenc
2	155.8	98.0	10442	22	AAF24702 Nucleotide sequenc
3	155.8	98.0	10474	22	AAF24685 Nucleotide sequenc
4	155.8	98.0	10474	22	AAF24686 Nucleotide sequenc
5	155.8	98.0	10474	22	AAF24707 Nucleotide sequenc
6	155.8	98.0	10474	22	AAF24708 Nucleotide sequenc
7	153.2	96.4	183999	22	AAF52831 Human ABC1 genomic
8	142.2	89.4	446	22	AAS04035 Partial human ABC1
9	142.2	89.4	7086	22	ABA09200 Human ABCA1 homolo

10	142.2	89.4	7086	22	AAK52667 Human polynucleoti
11	142.2	89.4	7260	22	AAD21326 Human ATP binding
12	142.2	89.4	7260	22	AAI70315 Human ATP binding
13	142.2	89.4	7281	22	AAK51683 Human polynucleoti
14	142.2	89.4	9741	22	AAS06120 Human ABC1 DNA seq
15	142.2	89.4	9854	22	AAK506121 Human ABC1 DNA seq
16	142.2	89.4	10545	21	AAK69132 Human ABC1 gene ex
17	140.6	88.4	736	22	AAH07432 Human cDNA clone (
18	140.6	88.4	1556	22	AAH18606 Human cDNA sequenc
19	135.2	85.0	7857	21	AAK69388 Human ABC1 cholest
20	135.2	85.0	7860	22	AAK69386 Human ABC1 cholest
21	135.2	85.0	7860	22	AAK69387 Human ABC1 cholest
22	135.2	85.0	7861	21	AAK69385 Human ABC1 cholest
23	135.2	85.0	7864	21	AAK69386 Human ABC1 cholest
24	135.2	85.0	7864	21	AAK69387 Human ABC1 cholest
25	135.2	85.0	7864	21	AAK69388 Human ABC1 cholest
26	135.2	85.0	7864	21	AAK69389 Human ABC1 cholest
27	35.8	22.5	50885	22	AAK70336 Human immune/haema
28	35.4	22.3	37	22	AAK93084 ABC1 polymorphism
29	35	22.0	298	21	AAK06182 Human secreted pro
30	35	22.0	23024	22	AAK25499 Nucleotide sequenc
31	34.6	21.8	534	21	AAK75311 Human ORFX ORF866
32	34.6	21.8	1133	22	AAD05589 Human secreted pro
33	33.2	20.9	5669	22	AAK08707 Human PD-ATP-bind
34	33.2	20.9	6522	22	AAK08706 Human PD-ATP-bind
35	31.4	19.7	6607	22	AAK54812 Nucleotide sequenc
36	31.4	19.7	8217	22	AAK07164 Human reproductive
37	30.4	19.1	485	21	AAK01004 Human secreted pro
38	30.4	19.1	532	21	AAK01005 Human secreted pro
39	30.4	19.1	2043	18	AAK79627 Human Doc2-beta ge
40	30.2	19.0	1093	22	ABA08699 Human ESRPI protei
41	30.2	19.0	1727	21	AAK66225 Human cDNA encodin
42	30.2	19.0	2143	22	AAI59934 Human polynucleoti
43	30.2	19.0	2644	22	AAK77695 Human Traf4 bindin
44	30.2	19.0	2668	22	AAK44707 Novel protein kina
45	30.2	19.0	2853	21	AAK21857 Human breast and o

ALIGNMENTS

RESULT 1

AAF24680

ID AAF24680 standard; DNA; 10442 BP.

AC AAF24680;

XX 20-APR-2001 (first entry)

DT Nucleotide sequence of a human ABC1 polypeptide.

DE Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

PH CDS 291..7076

FT /*tag= a

FT /product= "ABC1 polypeptide"

XX WO200078972-A2.

PN 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

PR 14-SEP-1999; 99US-0153872.

PR 19-NOV-1999; 99US-0166573.

CC genes are useful for developing screening assays to screen for compounds
CC that regulate the expression of genes associated with cholesterol
CC transport. The genes and proteins are also useful for are also useful
CC as diagnostic indicators of cardiovascular disease and other disorders
CC associated with hypercholesterolemia.

XX SQ Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

Query Match 98.0%; Score 155.8; DB 22; Length 10474;
Best Local Similarity 98.7%; Pred. No. 1.7e-36;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggcgctccctgtgtcagctctgagcgtgctccagggctccc 60
|||||

Db 230 ttaatgaccagccacggcgctccctgtgtcagctctgagcgtgctccagggctccc 289
|||||

QY 61 gagccacagctggcgctggtgaggaacatggcatgttgccctcagctgaggttg 120
|||||

Db 290 gagccacagctggcgctggtgaggaacatggcatgttgccctcagctgaggttg 349
|||||

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
|||||

Db 350 ctgctgtggaagaacctcactttcagaagaagacaaaca 388
|||||

RESULT 7

AAF92831
ID AAF92831 standard; DNA; 183999 BP.

XX AC AAF92831;

DT 17-MAY-2001 (first entry)

DE Human ABC1 genomic DNA.

XX KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.

XX OS Homo sapiens.

XX PN WO200115676-A2.

XX PD 08-MAR-2001.

XX PF 01-SEP-2000; 2000WO-IB01492.

XX PR 01-SEP-1999; 99US-0151977.

XX PR 15-MAR-2000; 2000US-0526193.

XX PR 23-JUN-2000; 2000US-0213958.

XX PA (UYBR-) UNIV BRITISH COLUMBIA.

XX PA (XENO-) XENON GENETICS INC.

XX PI Hayden MR, Brooks-Wilson AR, Pimstone SM, Clee SM;

XX DR WPI; 2001-244356/25.

XX PT Treating a lower than normal high density lipoprotein-cholesterol
(HDL-C) level, a higher than normal triglyceride level, or a
XX PT cardiovascular disease, by administering a compound that modulates LXR-
XX PT or RXR-mediated transcriptional activity -

XX PS Claim 8; Fig 1; 317pp; English.

XX CC The present invention relates to a method for treating a patient
XX CC diagnosed as having a lower than normal high density
XX CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
XX CC triglyceride level, or a cardiovascular disease, involving
XX CC administering a compound that modulates LXR- or RXR-mediated
XX CC transcriptional activity or ABC1 expression or activity.
XX CC The LXR gene product may be used in an assay to identify
XX CC compounds useful for the treatment of a disease or condition selected a
XX CC lower than normal HDL cholesterol level, a higher than normal

CC triglyceride level, and a cardiovascular disease.

XX SQ Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 96.4%; Score 153.2; DB 22; Length 183999;
Best Local Similarity 97.5%; Pred. No. 2.1e-35;
Matches 155; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggcgctccctgtgtcagctctgagcgtgctccagggctccc 60
|||||

Db 53229 ttaatgaccagccacggcgctccctgtgtcagctctgagcgtgctccagggctccc 53288
|||||

QY 61 gagccacagctggcgctggtgaggaacatggcatgttgccctcagctgaggttg 120
|||||

Db 53289 gagccacagctggcgctggtgaggaacatggcatgttgccctcagctgaggttg 53348
|||||

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
|||||

Db 53349 ctgctgtggaagaacctcactttcagaagaagacaaaca 53387
|||||

RESULT 8

AAS04035

ID AAS04035 standard; cDNA; 446 BP.

XX AC AAS04035;

XX DT 12-SEP-2001 (first entry)

XX DE Partial human ABC1 cDNA sequence.

XX KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;

XX KW cardiovascular; neurological; Tangier disease; LCAT deficiency;

XX KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 185..438

XX FT /*tag= a

XX FT /product= "Human ABC1 protein, amino acids 1 to 60"

XX PN WO200130848-A2.

XX PD 03-MAY-2001.

XX PF 26-OCT-2000; 2000WO-EP10886.

XX PR 26-OCT-1999; 99EP-0402668.

XX PR 01-MAR-2000; 2000US-0186260.

XX PA (AVET) AVENTIS PHARMA SA.

XX PI Denefle P, Rosier-Montus M, Arnould-Requigne I, Prades C, Naudin L;

XX PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;

XX PI Dean M;

XX DR WPI; 2001-316327/33.

XX DR P-PSDB; AAU02176.

XX PT New human ABC1 nucleic acids and peptides for treating

XX PT atherosclerosis, malaria and diabetes -

XX PS Example 2; Page 167; 368pp; English.

XX CC The sequence represents the partial coding sequence of human ABC1,

XX CC which encodes amino acids 1-60 of the human ABC1 protein. The nucleic

XX CC acid sequence, primers and probes derived from the ABC1 sequence, and

XX CC polypeptides and vectors are useful for the prevention of

XX CC atherosclerosis, in a subject affected by a dysfunction in the reverse

XX CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is

XX CC useful for screening for an active ingredient for the prevention or

CC treatment of a disease resulting from dysfunction in the reverse
 CC transport of cholesterol. The nucleic acids and polypeptides are also
 CC useful for treating and preventing cardiovascular and neurological
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-
 CC cholesterol (LCAT) deficiency, malaria and diabetes.
 XX
 SQ Sequence 446 BP; 96 A; 123 C; 112 G; 115 T; 0 other;

Query Match 89.4%; Score 142.2; DB 22; Length 446;
 Best Local Similarity 97.5%; Pred. No. 8.5e-33;
 Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 ttaatgacacgacggcgctcctgctgctcagctgcccgtgctccagggtccc 60
 |||||
 Db 93 ttaatgacacgacac-gggctcctgctgctgagctgcccgtgctccagggtccc 151
 |||||

Qy 61 gagccacacgtggcgctgctgctgaggaacatggcatgttgccctcagctgaggttg 120
 |||||
 Db 152 gagccacacgtggcgctgctgctgaggaacatggctgttgccctcagctgaggttg 211
 |||||

Qy 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 Db 212 ctgctgtggaagaacctcactttcagaagaagacaaaca 250
 |||||

RESULT 9
 ABA09200
 ID ABA09200 standard; cDNA; 7086 BP.
 AC ABA09200;
 XX
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human ABCA1 homologue-encoding cDNA, SEQ ID NO:976.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytotatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI; 2001-457740/49.
 DR
 DR P-PSDB; ABB11956.
 XX
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer
 XX
 XX Claim 1; Page 833-835; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, and hence
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth factor activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease, and
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.
 XX
 SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 89.4%; Score 142.2; DB 22; Length 7086;
 Best Local Similarity 97.5%; Pred. No. 1.7e-32;
 Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 ttaatgacacgacggcgctcctgctgctcagctgcccgtgctccagggtccc 60
 |||||
 Db 212 ttaatgacacgacac-ggctcctgctgctgagctgcccgtgctccagggtccc 270
 |||||

Qy 61 gagccacacgtggcgctgctgctgaggaacatggcatgttgccctcagctgaggttg 120
 |||||
 Db 271 gagccacacgtggcgctgctgctgaggaacatggctgttgccctcagctgaggttg 330
 |||||

Qy 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 Db 331 ctgctgtggaagaacctcactttcagaagaagacaaaca 369
 |||||

RESULT 10
 AAK52667
 ID AAK52667 standard; cDNA; 7086 BP.
 XX
 AC AAK52667;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 2196.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

AC AAI70315;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Human ATP binding cassette transporter 1 (ABCI) cDNA.
 XX
 KW ATP binding cassette transporter 1; ABC1; human; lipid disorder;
 KW cholesterol; cardiovascular disease; inflammatory disease;
 KW antiinflammatory; antilipemic; antiprosclerotic; dermatological;
 KW Tangier disease; coronary heart disease; diagnosis; gene therapy;
 KW polymorphism; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 321..7106
 FT FT /*tag= a
 FT CDS 501..7106
 FT FT /*tag= b
 FT FT /*note= "alternative open reading frame of AAI70314"
 FT FT replace(976,A)
 FT FT /*tag= c
 FT FT replace(1516,C)
 FT FT /*tag= d
 FT FT replace(2969,G)
 FT FT /*tag= e
 FT FT replace(3836,C)
 FT FT /*tag= f
 XX
 PN EPI136554-A1.
 XX
 XX 26-SEP-2001.
 XX
 XX 24-MAR-2000; 2000EP-0106401.
 XX
 XX 24-MAR-2000; 2000EP-0106401.
 XX (FARB) BAYER AG.
 XX
 PI Schmitz G, Bodzioch M;
 XX
 DR WPI; 2001-640389/74.
 DR P-PSDB; AAM50228.
 XX
 PT New adenosine triphosphate binding cassette transporter gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 PT cardiovascular diseases and inflammatory diseases
 XX
 PS Disclosure; Page 26-28; 41pp; English.
 XX
 CC The present sequence is that of cDNA encoding the human adenosine
 CC triphosphate (ATP) binding cassette transporter 1 (ABCI) protein
 CC (see AAM50227). The sequence includes an extended open reading
 CC frame (ORF) to that provided by the sequence in AAI70314, using
 CC an alternative ATG codon as initiation codon and thereby adding an
 CC extra 40 N-terminal amino acids to the encoded ABC1 protein (see
 CC AAM50228). The invention provides 4 common polymorphisms in the
 CC ABC1 gene. These were identified by sequencing the ABC1 gene in
 CC different Tangier kindreds. In the variant genes (numbering as in
 CC AAI70314), G is changed to A at position 596, T is changed to C at
 CC position 1136, A is changed to G at position 2589 or G is changed
 CC to C at position 3456, or any combination of these. All of these
 CC polymorphisms alter the amino acid sequence of ABC1 and therefore
 CC may affect its function. The 2 most common polymorphisms (G596A
 CC and A2589G) are both associated with a decreased in vitro ApoA-I
 CC mediated efflux of cholesterol from mononuclear phagocytes, a
 CC feature typical of Tangier disease. 3 Of the variants (G596A,
 CC A2589G and G3456C) are significantly increased in a population of
 CC men having low high density lipoprotein-cholesterol levels and
 CC established coronary heart disease (CHD) relative to CHD-free
 CC control subjects. The use of the provided ABC1 polymorphisms for
 CC the diagnosis and treatment of lipid disorders, cardiovascular
 CC diseases, and inflammatory diseases (e.g. psoriasis, lupus

CC erythematoses) is claimed. Modulation of ABC1 transcripts or
 CC proteins by antisense or ribozyme technology or RNA decoys is also
 CC claimed.
 XX
 SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;
 XX
 Query Match 89.4%; Score 142.2; DB 22; Length 7260;
 Best Local Similarity 97.5%; Pred. No. 1.7e-32;
 Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1 ttaattgaccacgacggcggtccctgctgtgcagctctgtgcgctgcctccagggttgc 60
 |||||
 DB 229 ttaattgaccacgacggcggtccctgctgtgcagctctgtgcgctgcctccagggttgc 287
 QY 61 gagccacacgctggcggtgcgtgcggtgcggtgcggtgcggtgcggtgcggtgcggtg 120
 |||||
 DB 288 gagccacacgctggcggtgcgtgcggtgcggtgcggtgcggtgcggtgcggtgcggtg 347
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 DB 348 ctgctgtggaagaacctcactttcagaagaagacaaaca 386
 RESULT 13
 AAK51683
 ID AAK51683 standard; cDNA: 7281 BP.
 XX
 AC AAK51683;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 228.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM78550.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy
 XX
 PS Claim 1; Page 1086-1096; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 89.4%; Score 142.2; DB 22; Length 7281;
 Best Local Similarity 97.5%; Pred. No. 1.7e-32;
 Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1 ttaatgaccagcagcggtccctgctgtcagctgtgagcgtggtcaggtctccc 60
 |||||
 Db 250 ttaatgaccagccac-ggcgtccctgctgtgagctgtggtcgtcgtcaggtctccc 308
 |||||
 QY 61 gagccacacgctggcgctgctggtgagggagacatggcatgttggtcctcagctgaggttg 120
 |||||
 Db 309 gagccacacgctgggggtgctggtgagggagacatggctgtgttggtcctcagctgaggttg 368
 |||||
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 Db 369 ctgctgtggaagaacctcactttcagaagaagacaaaca 407
 |||||

RESULT 14
 AAS06120
 ID AAS06120 standard; cDNA; 9741 BP.
 XX
 AC AAS06120;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human ABC1 DNA sequence #1.
 XX
 KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
 KW cardiovascular; neurological; Tangier disease; LCAT deficiency;
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 185..6967
 FT /*tag= a
 FT /product= "Human ABC1 protein"

WO200130848-A2.
 03-MAY-2001.
 26-OCT-2000; 2000WO-EPI0886.
 26-OCT-1999; 99EP-0402668.
 01-MAR-2000; 2000US-0186260.
 (AVET) AVENTIS PHARMA SA.
 Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;
 Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;
 Dean M;
 WPI: 2001-316327/33.
 P-PSDB; AAU02176.
 New human ABC1 nucleic acids and polypeptides for treating
 atherosclerosis, malaria and diabetes -

XX
 PS Claim 1; Page 204-208; 368pp; English.
 XX
 CC The sequence represents the coding sequence #1 of human ABC1. The
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,
 CC and polypeptides and vectors are useful for the prevention of
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is
 CC useful for screening for an active ingredient for the prevention or
 CC treatment of a disease resulting from dysfunction in the reverse
 CC transport of cholesterol. The nucleic acids and polypeptides are also
 CC useful for treating and preventing cardiovascular and neurological
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-
 CC cholesterol (LCAT) deficiency, malaria and diabetes.
 XX
 SQ Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;

Query Match 89.4%; Score 142.2; DB 22; Length 9741;
 Best Local Similarity 97.5%; Pred. No. 1.8e-32;
 Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1 ttaatgaccagcagcggtccctgctgtcagctgtggtcgtcgtcaggtctccc 60
 |||||
 Db 93 ttaatgaccagccac-ggcgtccctgctgtgagctgtggtcgtcgtcaggtctccc 151
 |||||
 QY 61 gagccacacgctggcgctgctggtgagggagacatggcatgttggtcctcagctgaggttg 120
 |||||
 Db 152 gagccacacgctgggggtgctggtgagggagacatggctgtgttggtcctcagctgaggttg 211
 |||||
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
 |||||
 Db 212 ctgctgtggaagaacctcactttcagaagaagacaaaca 250
 |||||

RESULT 15
 AAS06121
 ID AAS06121 standard; cDNA; 9854 BP.
 XX
 AC AAS06121;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human ABC1 DNA sequence #2.
 XX
 KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
 KW cardiovascular; neurological; Tangier disease; LCAT deficiency;
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 298..7078
 FT /*tag= a
 FT /product= "Human ABC1 protein"

WO200130848-A2.
 03-MAY-2001.
 26-OCT-2000; 2000WO-EPI0886.
 26-OCT-1999; 99EP-0402668.
 01-MAR-2000; 2000US-0186260.
 (AVET) AVENTIS PHARMA SA.
 Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;
 Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;
 Dean M;
 WPI: 2001-316327/33.
 P-PSDB; AAU02176.

```

XX New human ABC1 nucleic acids and polypeptides for treating
PT atherosclerosis, malaria and diabetes
XX
XX Claim 1; Page 209-213; 368pp; English.
XX
XX The sequence represents the coding sequence #2 of human ABC1. The
XX nucleic acid sequence, primers and probes derived from the ABC1 sequence,
XX and polypeptides and vectors are useful for the prevention of
XX atherosclerosis, in a subject affected by a dysfunction in the reverse
XX transport of cholesterol. The polypeptide encoded by the ABC1 gene is
XX useful for screening for an active ingredient for the prevention or
XX treatment of a disease resulting from dysfunction in the reverse
XX transport of cholesterol. The nucleic acids and polypeptides are also
XX useful for treating and preventing cardiovascular and neurological
XX pathologies, and other diseases e.g. Tangier disease, lecithin-
XX cholesterol (LCAT) deficiency, malaria and diabetes.
XX
SQ Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;

Query Match      89.4%; Score 142.2; DB 22; Length 9854;
Best Local Similarity 97.5%; Pred. NO. 1.8e-32;
Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 ttaatgaccagccagggcgctcctgtcagctctgcccgtgcttcacagggtccc 60
   |||||||
Db 206 ttaatgaccagccac-ggcgtccctgtgtgagctctgcccgtgcttcacagggtccc 264
   |||||||

Qy 61 gagccacacgtggcgctgtgctgagggagacatggcatgttggcctcagctgaggttg 120
   |||||||
Db 265 gagccacacgtgggggtgctgctgagggagacatggttggcctcagctgaggttg 324
   |||||||

Qy 121 ctgctgtggaagacacctcactttcagaagaagacaaaca 159
   |||||||
Db 325 ctgctgtggaagacacctcactttcagaagaagacaaaca 363
   |||||||

```

Search completed: September 20, 2002, 03:09:36
Job time: 10645 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: September 20, 2002, 03:21:58 ; Search time 110.16 Seconds
(without alignments)
354.536 Million cell updates/sec

Title: US-09-846-456-5
Perfect score: 159
Sequence: 1 ttaatgaccagcagcgcg.....cttcagaagaagacaaaca 159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCRU.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	18.9	4771	2	US-08-866-650-2
2	30	18.9	4771	2	US-09-021-287-2
3	30	18.9	4771	4	US-09-240-473-2
4	30	18.9	33529	4	US-09-144-085-3
5	29.4	18.5	3147	2	US-08-781-802-7
6	29.4	18.5	3147	4	US-08-694-078-7
7	29.4	18.5	3147	4	US-09-058-260-7
8	28.6	18.0	2861	4	US-08-482-073-10
9	28.6	18.0	3647	1	US-07-914-281-7
10	28.6	18.0	3647	1	US-08-393-246-7
11	28.6	18.0	3647	2	US-08-525-058A-7
12	28.6	18.0	3647	2	US-08-696-731-7
13	28.6	18.0	3647	4	US-09-042-531-7
14	28	17.6	1146	1	US-08-482-385A-1
15	28	17.6	2728	1	US-08-482-385A-5
16	27.8	17.5	49136	4	US-09-422-869-1
17	27.6	17.4	318	2	US-08-646-981-4
18	27.6	17.4	1257	3	US-08-640-906-1
19	27.6	17.4	1257	4	US-09-395-936-1
20	27.6	17.4	2001	4	US-09-422-869-7
21	27.4	17.2	733	4	US-09-082-092-15
22	27.4	17.2	6176	3	US-08-911-321-6
23	27.4	17.2	4411529	4	US-09-103-840A-1
24	27.2	17.1	750	3	US-09-010-809-12
25	27.2	17.1	1633	4	US-09-119-788-1
26	27.2	17.1	1864	2	US-08-673-388-9
27	27.2	17.1	1864	2	US-08-614-877-9

c	28	27.2	17.1	3300	2	US-08-928-692-29	Sequence 29, Appl
c	29	27	17.0	3647	5	PCT-US91-00899-4	Sequence 4, Appl
c	30	26.6	16.7	1238	2	US-08-330-290-11	Sequence 11, Appl
	31	26.4	16.6	884	2	US-08-901-200A-11	Sequence 11, Appl
	32	26.4	16.6	884	3	US-09-219-391-11	Sequence 11, Appl
	33	26.4	16.6	1723	1	US-07-841-646-28	Sequence 28, Appl
	34	26.4	16.6	1723	1	US-07-901-703-10	Sequence 10, Appl
	35	26.4	16.6	1723	1	US-08-147-023-28	Sequence 28, Appl
	36	26.4	16.6	1723	1	US-08-206-864-3	Sequence 3, Appl
	37	26.4	16.6	1723	1	US-08-278-729A-20	Sequence 20, Appl
	38	26.4	16.6	1723	1	US-08-480-528A-7	Sequence 7, Appl
	39	26.4	16.6	1723	1	US-08-479-666-7	Sequence 7, Appl
	40	26.4	16.6	1723	1	US-08-155-343A-20	Sequence 20, Appl
	41	26.4	16.6	1723	1	US-08-406-672-20	Sequence 20, Appl
	42	26.4	16.6	1723	1	US-08-643-563A-20	Sequence 20, Appl
	43	26.4	16.6	1723	1	US-08-447-570-28	Sequence 28, Appl
	44	26.4	16.6	1723	1	US-08-643-763A-20	Sequence 20, Appl
	45	26.4	16.6	1723	1	US-08-462-623-20	Sequence 20, Appl

ALIGNMENTS

```

RESULT 1
US-08-866-650-2
; Sequence 2, Application US/08866650
; Patent No. 5939321
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,650
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mtll protein"
US-08-866-650-2

```

Query Match 18.9%; Score 30; DB 2; Length 4771;
Best Local Similarity 52.4%; Pred. No. 3.2;
Matches 66; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

RESULT

US-09-240-473-2
; Sequence 2, Application US/09240473
; Patent No. 6297011

; GENERAL INFORMATION:

APPLICANT: Takanara, Kazuniko
 ADDRESS: Hoffman, Guy C

NUMBER OF SEQUENCES: 13

ADDRESSEE: Quarles & Brady

CITY: Madison

COUNTRY: US
ZIP: 53703

MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA:

CLASSIFICATION.

NAME: Berson, Bennett J

REFERENCE/DOCKET NUMBER: 9602

TELEPHONE: 608-251-3000
TELETYPE: 608-251-0166

: SEQUENCE CHARACTERISTICS:

; TYPE: nucleic acid

```
; TOPOLOGY: Linear
```

```

;
; FEATURE:
; NAME/KEY: CDS

```

```

;
; OTHER INFORMATION:  /product=

```

Query Match

PM101500, CONSERVATIVE 0,

10

201 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817

0v 123 actata 128

Db 261 GCTGAG 266

;
CURRENT APPL
;
APPLICAT

```

; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704
; FILING DATE: 10-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,995
; FILING DATE: 07-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 245..1231
; OTHER INFORMATION: /note= "TspA E101 sequence longest
; open reading frame; other possible start codons are TTG/leu9;
; TTG/leu13; TTG/leu15; GTG/val43"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 245..1231
; US-08-694-078-7

Query Match 18.5%; Score 29.4; DB 4; Length 3147;
Best Local Similarity 56.8%; Pred. No. 4.4;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 45 gcttcacgtggtgctgcgtgaggaacactca 139
|||||
Db 2798 GCCTCCCGGACTACCTCCCGCGGCTTTAGGGCTCTAGTACCACCCATCCTG 2857
|||||

QY 105 gctcacgtgaggtgctgctgtggaagaactca 139
|||||
Db 2858 GCGTACGCCAGGATGGGGCCCGGTAAGGCCTTA 2892
|||||

RESULT 7
US-09-058-260-7
; Sequence 7, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; CURRENT FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08

; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704
; FILING DATE: 10-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,995
; FILING DATE: 07-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 245..1231
; OTHER INFORMATION: /note= "TspA E101 sequence longest
; open reading frame; other possible start codons are TTG/leu9;
; TTG/leu13; TTG/leu15; GTG/val43"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 245..1231
; US-08-694-078-7

Query Match 18.5%; Score 29.4; DB 4; Length 3147;
Best Local Similarity 56.8%; Pred. No. 4.4;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 45 gcttcacgtggtgctgcgtgaggaacactca 139
|||||
Db 2798 GCCTCCCGGACTACCTCCCGCGGCTTTAGGGCTCTAGTACCACCCATCCTG 2857
|||||

QY 105 gctcacgtgaggtgctgctgtggaagaactca 139
|||||
Db 2858 GCGTACGCCAGGATGGGGCCCGGTAAGGCCTTA 2892
|||||

RESULT 8
US-08-482-073-10/C
; Sequence 10, Application US/08482073
; Patent No. 6307025
; GENERAL INFORMATION:
; APPLICANT: Hession, Catherine A.
; APPLICANT: Lobb, Roy R.
; APPLICANT: Goelz, Susan E.
; APPLICANT: Osborn, Laurelee
; APPLICANT: Benjamin, Christopher D.
; APPLICANT: Rosa, Margaret D.
; TITLE OF INVENTION: ENDOTHELIAL CELL-LEUKOCYTE ADHESION
; TITLE OF INVENTION: MOLECULES (ELAMS) AND MOLECULES INVOLVED IN LEUKOCYTE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,073
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,336
; FILING DATE:
; APPLICATION NUMBER: US 07/608298
; FILING DATE: 31-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US 90/02357
; FILING DATE: 27-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/452675

```

; TELEX: 248855 OPAT UR


```

; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,531
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; US-09-042-531-7

```

```

Query Match 18.0%; Score 28.6; DB 4; Length 3647;
Best Local Similarity 57.1%; Pred. No. 8;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 5 tgaccagccagggcgtccctgtgtcagctgtgcccgttcgcttcacagggctcccgagc 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1733 TGAGCTGGCCAGGACGCCAACCGCGGACCGCCGCTCCCTTTCTTCCACTGCGCCGGG 1674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 65 cacacgctggcgctgctggctgagggacat 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1673 CCCAGCGGCCCCGACGAGGCCCCGGGAGCCT 1643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-08-482-385A-1/c
; Sequence 1, Application US/08482385A
; Patent No. 5728561
; GENERAL INFORMATION:
; APPLICANT: DENOYA,, CLAUDIO D.
; TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA
; KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: PETER C. RICHARDSON
; STREET: 235 EAST 42ND STREET, 20TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A
; ZIP: 10017-5755
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,385A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SHEYKA,, ROBERT F.
; REGISTRATION NUMBER: 31,304
; REFERENCE/DOCKET NUMBER: PC8346C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-573-1189
; TELEFAX: 212-573-1939
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-482-385A-1

```

```

Query Match 17.6%; Score 28; DB 1; Length 1146;
Best Local Similarity 55.0%; Pred. No. 8.7;
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 15 cggcgctccctgctgcagctgtgcgctgccttcagggtcccgagccacacgctgg 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 CGGCCCTGCTTGGTGAGAGCCGTGGCTGGTGTGTAGCGGGCGCGGCACCACTCC 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 75 gcgtgctggtgagggaacatggcattgttgccctcagctg 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 GCGTACAGCGCGGCAGCAGTAGCGGGTGGCTCCGCGG 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-08-482-385A-5/c
; Sequence 5, Application US/08482385A
; Patent No. 5728561
; GENERAL INFORMATION:
; APPLICANT: DENOYA,, CLAUDIO D.
; TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA
; KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER C. RICHARDSON
; STREET: 235 EAST 42ND STREET, 20TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A
; ZIP: 10017-5755
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,385A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SHEYKA,, ROBERT F.
; REGISTRATION NUMBER: 31,304

```

Fri Sep 20 08:03:46 2002

```

; REFERENCE/DOCKET NUMBER: PC8346C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-573-1189
; TELEFAX: 212-573-1939
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2728 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-482-385A-5

```

```

Query Match      17.6%; Score 28; DB 1; Length 2728;
Best Local Similarity 55.0%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 15 cgggcgtccctgctcagcttgccgctgacctccaggctcccgagccacacgctgg 74
Db 626 CGGCCCTGCTTGGTGAGAGCCGTGGCTGCTGTTGTAGCGCGCGCGGCACCACTCC 567

Qy 75 gcgtgctggctgaggggaacatggcatgttgccctcagctg 114
Db 566 GCGTACAGCGCGCGCAGCAGTAGCGGGTGGCTCGCGCGG 527

```

Search completed: September 20, 2002, 03:22:05
Job time: 14214 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:18 ; Search time 3895 Seconds
(without alignments)
550.967 Million cell updates/sec

Title: US-09-846-456-5
Perfect score: 159
Sequence: 1 ttaatgaccgcccacgggag.....ctttcagaagaagacaaca 159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	140.6	88.4	736	9	AU135588
2	72.8	45.8	619	9	BB657864
3	70.8	44.5	292	10	Z44377
4	70.8	44.5	535	10	BG384217
5	37	23.3	440	9	AA914462
6	36	22.6	533	9	AV845237
7	35.6	22.4	514	12	BH087817
8	35.4	22.3	482	12	AO848626
9	35	22.0	571	9	AV862022
10	34.8	21.9	301	9	AW751638
11	34.8	21.9	365	9	BB843149
12	34.8	21.9	516	10	BF484412
13	34.8	21.9	530	10	BE471178
14	34.6	21.8	344	10	BM149133
15	34.6	21.8	1574	10	BG325753
16	34.4	21.6	571	9	AV864011
17	34.4	21.6	579	9	AV892280

18	34	21.4	598	12	BH087727
19	33.8	21.3	668	10	BG398045
20	33	20.8	335	9	AW313960
21	33	20.8	477	12	AQ885360
22	33	20.8	575	10	BI541694
23	33	20.8	791	10	BG675383
24	33	20.8	1901	10	BF128237
25	32.8	20.6	543	10	BG747266
26	32.8	20.6	997	12	CNS006DN
27	32.6	20.5	653	9	BB132336
28	32.6	20.5	870	9	AL530607
29	32.4	20.4	658	9	BB866093
30	32.4	20.4	667	9	BB653232
31	32.2	20.3	684	10	BG700305
32	32.2	20.3	715	10	BG717361
33	32.2	20.3	971	10	BI915421
34	32	20.1	277	9	BB265873
35	32	20.1	650	9	AL636346
36	32	20.1	771	10	BI770317
37	32	20.1	777	10	BG830641
38	32	20.1	884	10	BI668321
39	32	20.1	935	9	AL539815
40	31.8	20.0	548	12	AQ452008
41	31.8	20.0	615	10	BE981503
42	31.8	20.0	682	9	BB644307
43	31.8	20.0	724	10	BE981359
44	31.8	20.0	801	12	AQ745776
45	31.6	19.9	472	9	AA726802

ALIGNMENTS

RESULT 1

AU135588
LOCUS AU135588
DEFINITION PLACEL Homo sapiens cDNA clone PLACE1002437 5', mRNA
ACCESSION AU135588
VERSION AU135588.1 GI:10996127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 736)
AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
Location/Qualifiers
1. .736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pME18SFL3"

BASE COUNT 163 a 199 c 199 g 170 t
ORIGIN 5 others

```

Query Match      88.4%; Score 140.6; DB 9; Length 736;
Best Local Similarity 96.9%; Pred. No. 1.4e-28;
Matches 154; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 ttaatgacagccagcggcgctcctcgtcgtcagctctggtccgctgcctccagggctccc 60
|||||
Db 222 TTAATGACAGCCAC-GGCGTCCCTGCTGTGAGCTCTGGCGCTGCCCTCCAGGCTCCC 280
|||||

QY 61 gagccacgcgtggcgctgctgaggaacatggcatgttggctcagctaggttg 120
|||||
Db 281 GAGCCACAGCGTGGGGTCTGGCTGAGGGAACATGGCTTGTGGCCCAAGCTAGGTTG 340
|||||

QY 121 ctgctgtggaagaacctcaacttcagagaagacaaaca 159
|||||
Db 341 CTGCTGTGGAAGAACCTCACCTTCAGAGAAGACAAACA 379
|||||

RESULT 2
LOCUS BB657864 619 bp mRNA linear EST 26-OCT-2001
DEFINITION BB657864 RIKEN full-length enriched, 12 days embryo eyeball Mus
musculus cDNA clone D230019D04 5', mRNA sequence.
ACCESSION BB657864
VERSION BB657864.1 GI:16491690
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 619)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,I., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```

```

e mouse tissues.
Location/Qualifiers
1. .619
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D230019D04"
/clone_lib="RIKEN full-length enriched, 12 days embryo
eyeball"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCCCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATTCGAGTTCTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 126 a 185 c 173 g 134 t 1 others
ORIGIN
Query Match 45.8%; Score 72.8; DB 9; Length 619;
Best Local Similarity 77.4%; Pred. No. 5.9e-10;
Matches 130; Conservative 0; Mismatches 27; Indels 11; Gaps 3;
QY 1 ttaatgacagccagcggcgctcctcgtcgtcagctctggtccgctcctccagggctccc 60
|||||
Db 222 TTAATGACAGCCAC-AGAGTCACAGCTGTGCTGTGCTGCT-CCCTCCAGGCTCTC 279
|||||

QY 61 gagccaca-----cgctggcgctgctgaggaacatggcatgttggctcag 111
|||||
Db 280 GAGCCGACAGCGAGTCTGCTGTGGTCCGCTGTGGTGCATGGCTGTGGGCTCTCAG 339
|||||

QY 112 ctgaggtgtcgtgtggaagaacctcactttcagaagaagacaaaca 159
|||||
Db 340 TTAAGGCTGCTGCTGTGGAAGAATCTGACATTTTCGAGGAGAGACAAACA 387
|||||

RESULT 3
LOCUS 244377
DEFINITION HSC12B081 normalized infant brain cDNA Homo sapiens cDNA clone
c-12b08, mRNA sequence.
ACCESSION 244377
VERSION 244377.1 GI:573506
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 292)
Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
Sebastiani-Kabaktchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800

```


	Query Match	44.5%	Score 70.8;	DB 10;	Length 535;
	Best Local Similarity	79.2%;	Pred. No. 2e-09;		
	Matches	84;	Conservative	0;	Mismatches 22; Indels 0; Gaps
QY	54	ggctcccgagccacagctggcgctgctgaaggaacatggcatgttggcctcagct	113		
Db	244	GGACCCACAGCAGCGTGGTGGTGCACGCTGAGGTACATGCTTATTGGACTCACT	303		
QY	114	gaggttgcctgctggaagaaccttcacttcagaagaagacaaca	159		
Db	304	GAGGTTACTGCTGTGGAGAACCTCACTTTTCAGAAGAAGACAAACA	349		

RESULT	5
AA914462/c	
LOCUS	
DEFINITION	AA914462 vz01f08.r1 Soares_mammary_gland_NBMG Mus musculus CDNA clone IMAGE:3114471 5' mRNA sequence.
ACCESSION	AA914462
VERSION	AA914462.1 GI:3053854
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus Eukaryota; Metazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

[illegible]

RESULT	9
AV862022/c	
LOCUS	
DEFINITION	

```

/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
80 a 70 c 85 g .66 t
BASE COUNT

```

FEATURES	
SOURCE	Location/Qualifiers
e mouse tissues.	1. 365
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="F43009P04"
	/clone_lib="RIKEN full-length enriched, 6 days neonate spleen"
	/tissue_type="spleen"
	/dev_stage="6 days neonate"
BASE COUNT	80 a 114 c 96 g 75 t
ORIGIN	
Query Match	21.9%; Score 34.8; DB 9; Length 365;
Best Local Similarity	62.8%; Pred. No. 14;
Matches 54; Conservative	0; Mismatches 32; Indels 0; Gaps 0;

Db	78	GGGCGCTCAGACGAGTCATCATGCGCTTAGGCACACAGCTGATGCTTCTGCTGTGAAA	137
QY	133	aaactcactttcagaagaagacaaac	158
Db	138	AATTACACCTATCGACGGAGACAACC	163
RESULT	12		
BF484412			
LOCUS		516 bp	linear
DEFINITION		WHE2323_B07_D132S wheat pre-anthesis spike cDNA library	EST 06-DEC-2000
ACCESSION		aestivum cDNA clone WHE2323_B07_D13,	Triticum
VERSION		BF484412	sequence.
KEYWORDS		BF484412.1	GI:11567713
SOURCE		EST	
ORGANISM		bread wheat.	
		Triticum aestivum	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;	
		1 (bases 1 to 516)	
REFERENCE		Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han	
AUTHORS		P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,	
		Seaton, C.L. and Tong, J.C.	
TITLE		The structure and function of the expressed portion of the wheat	
JOURNAL		genomes - pre-anthesis spike cDNA library	
COMMENT		Unpublished (2000)	
		Contact: Olin Anderson	

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

Email: oanderns@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES

source

Location/Qualifiers

```
1. .516
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE2323_B07_D13"
/clone_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
phagemids in the TJ Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
```

BASE COUNT

ORIGIN

106 a 177 c 136 g 97 t

Query Match

Best Local Similarity 21.9%; Score 34.8; DB 10; Length 516;

Matches 63; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 19 cgtccctgtcagctctgcccgtctccagggctccagggccgcacacgtggcggt 78

Db 276 CGCCGCTGCCGACATCTCCGACCCATGACGACGAGCTCAGTCAAGACAGATGTCGGCGAG 335

QY 79 gctggtgagggaacatggtgctggtgctcagctcagctgaggtgtgtgtg 128

Db 336 GGGAGAAAGGAAGATGGGATGCCGATCATCTCTGGAGCGCGCTGTG 385

RESULT 13

BE471178

LOCUS

BE471178 530 bp mRNA linear EST 28-JUL-2000

DEFINITION WHE0285_G07_N13ZS Wheat drought-stressed seedling cDNA library

Triticum aestivum cDNA clone WHE0285_G07_N13, mRNA sequence.

ACCESSION BE471178

VERSION BE471178.1

KEYWORDS GI:9561669

SOURCE EST.

bread wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Triticum.

1 (bases 1 to 530)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han

, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,

Seaton, C.L. and Tong, J.C.

The structure and function of the expressed portion of the wheat

genomes - Drought-stressed seedling cDNA library

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderns@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES

source

Location/Qualifiers

```
1. .530
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0285_G07_N13"
/clone_lib="Wheat drought-stressed seedling cDNA library"
/tissue_type="Seedling without endosperm"
/dev_stage="Five day old seedling"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Five-day old
seedlings were incubated for one day at 90% RH. After
removing endosperm, seedlings were transferred to
desiccator jar containing saturated MgSO4 at room
temperature for 24 hr. The tissue, total RNA, and poly(A)
RNA were prepared, a cDNA library was made, and the cDNA
clones were in vivo excised to give pBluescript phagemids
in the TJ Close lab (Choi, Close, Fenton) at the
University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
```

BASE COUNT

ORIGIN

113 a 178 c 143 g 96 t

Query Match

Best Local Similarity 21.9%; Score 34.8; DB 10; Length 530;

Matches 63; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 19 cgtccctgtcagctctgcccgtctccagggctccagggccgcacacgtggcggt 78

Db 275 CGCCGCTGCCGACATCTCCGACCCATGACGACGAGCTCAGTCAAGACAGATGTCGGCGAG 334

QY 79 gctggtgagggaacatggtgctggtgctcagctcagctgaggtgtgtgtg 128

Db 335 GGGAGAAAGGAAGATGGGATGCCGATCATCTCTGGAGCGCGCTGTG 384

RESULT 14

BM149133

LOCUS

BM149133

DEFINITION

TCAAP2E6391

Baylor-HGSC project-TCAA

Homo sapiens

cDNA clone

TCAAP6391, mRNA

sequence.

ACCESSION

BM149133

VERSION

BM149133.1

KEYWORDS

GI:17170474

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa;

Chordata; Craniata;

Vertebrata; Euteleostomi;

Mammalia; Eutheria;

Primates; Catarrhini;

Homidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 344)

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,

Gunnaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project (2001)

Unpublished (2001)

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@txccc.org

Seq primer: M13 primer.

FEATURES

source

1. .344

Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCRAP6391"
/clone_lib="pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-PCAA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10B"
/notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer [5'GGAGGACTGAGCGCGCAGGAGGAG(T)VN 3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand was primed with a BamHI-dC primer [5'AGAGCTCGGATCGCGCGCATATATAAT(C) 3']. Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and SalI sites of lambda pSB vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Garnicki P, Westover A, Nishiyama Y, Ohsumi T, Itoh M, Nagaoaka S, Sasakini, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y, High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      57 a  108 c  106 g  73 t
ORIGIN

```

```

Query Match      21.8%; Score 34.6; DB 10; Length 344;
Best Local Similarity 70.8%; Pred. No. 16;
Matches 46; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY  91  aacatgcatgttgccctcagctgaggtgctgtggaagacctcactttcagaaga 150
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   232  ACATGCGCTTCGGACACAGCTGATGCTGCTCTGGAGAAATTCATGATCGCCGG 291

QY  151  agaca 155
      |||||
Db   292  AGACA 296

```

```

RESULT 15
BG325753
LOCUS      602424464F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562559 5',
DEFINITION mRNA sequence.
ACCESSION BG325753
VERSION    BG325753.1 GI:13132177
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1574)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICMI274 row: p column: 16
            High quality sequence start: 5
            High quality sequence stop: 178.
FEATURES
     source      1..1574
                /organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:4562559"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      305 a  588 c  432 g  249 t
ORIGIN

```

```

Query Match      21.8%; Score 34.6; DB 10; Length 1574;
Best Local Similarity 55.4%; Pred. No. 25;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

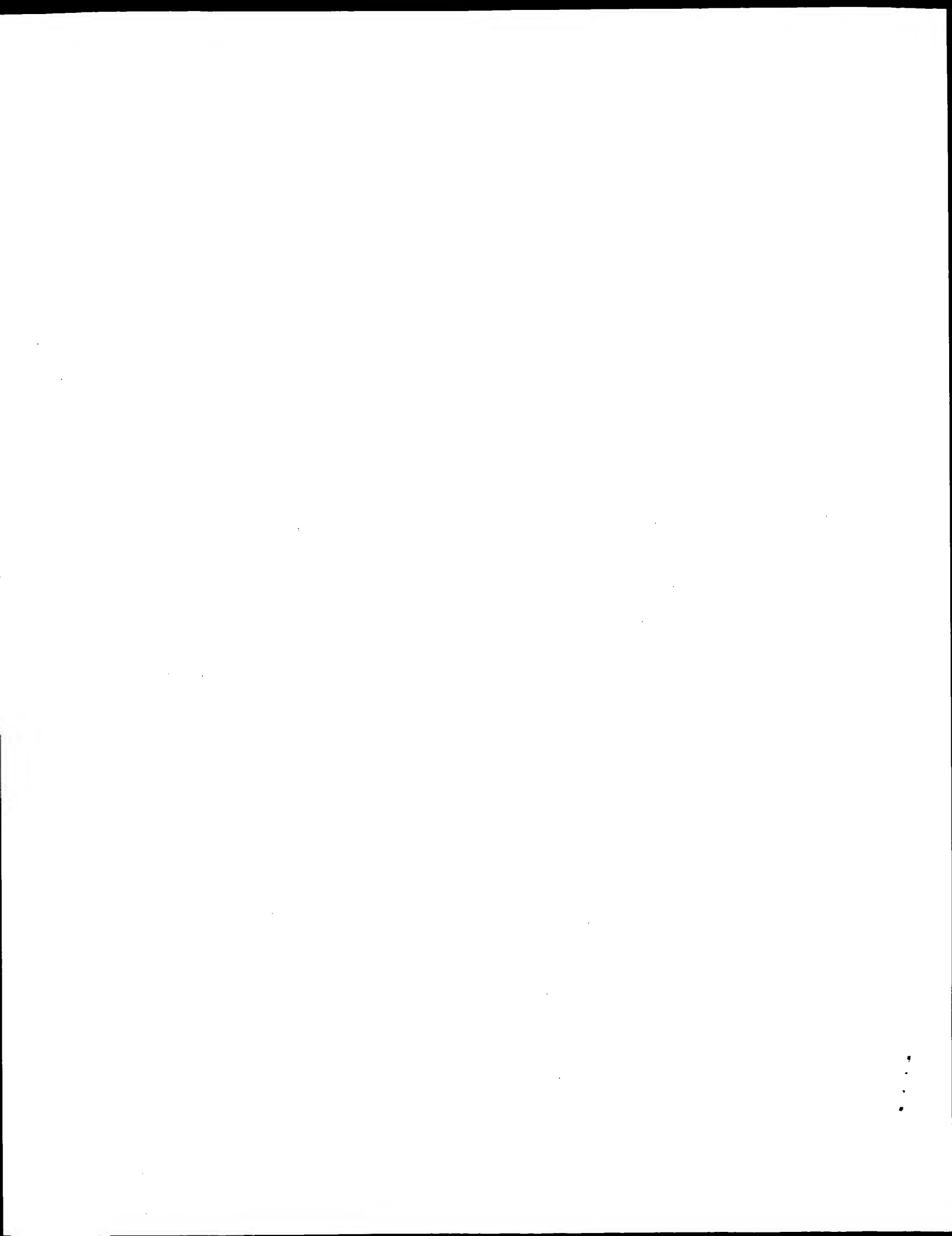
QY  2   taatgaccagccacggcgctccctgctgcagctctggcgccttcacagggtcccg 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1216  TCACGAGCCCCCCTTGTTCCTCCGCTGTCGCTTTCCTGCTGTCGCTGTCGCGG 1275

QY  62   agccacacgctggcgctgctgctgaggggaacatggcatgttgccctcagctgaggttgc 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1276  AGCCAAAGCGCACGGAGCGGGCTCTGCGAGTAGGAGAGAGCTTGGCCGCCCGGGGTGGAGC 1335

QY  122  t 122
Db   1336  T 1336

```

Search completed: September 20, 2002, 01:23:20
Job time: 7380 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:26 : Search time 5250.46 Seconds
(without alignments)
12877.660 Million cell updates/sec

Title: US-09-846-456-1
Perfect score: 3231
Sequence: 1 acaggcgatgtggcagtg.....gcccacatccccaccactt 3231

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 1797656 seqs, 10463268293 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hlg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description

1	3231	100.0	3231	6	AX351029	Sequence
2	2893	89.5	2893	6	AX351031	Sequence
3	1749	54.1	201144	9	AF287262	Homo sapi
4	1487	46.0	149034	9	AF275948	Homo sapi
5	891	27.6	96717	9	AL359182	Human DNA
6	814	25.2	183999	6	AX092589	Sequence
7	686	21.2	175064	2	AC012230	Homo sapi
8	589	18.2	1167	9	AF258623	Homo sapi
9	589	18.2	1167	9	AF258623S1	Sequence
10	541	16.7	1643	6	AX060715	Sequence
11	541	16.7	1643	6	AX060894	Sequence
12	348	10.8	69570	2	AC021246	Homo sapi
13	344	10.6	175064	2	AC012230	Homo sapi
14	336	10.4	1750	9	AF258627	Homo sapi
15	303	9.4	697	9	AF258627	Homo sapi
16	282	8.7	90598	2	AC021345	Homo sapi
17	228	7.1	7260	6	AX253452	Sequence
18	221	6.8	221	6	AX351032	Sequence
19	217	6.7	1556	9	AK024328	Homo sapi
20	205	6.3	9854	6	AX127831	Sequence
21	205	6.3	9854	6	AX139818	Sequence
22	197	6.1	10442	6	AX060713	Sequence
23	197	6.1	10442	6	AX060892	Sequence
24	197	6.1	10442	9	AF285167	Homo sapi
25	188	5.8	10474	6	AX060719	Sequence
26	188	5.8	10474	6	AX060721	Sequence
27	188	5.8	10474	6	AX060898	Sequence
28	188	5.8	10474	6	AX060900	Sequence
29	152	4.7	90698	2	AC021345	Homo sapi
30	117	3.6	117	6	AX351034	Sequence
31	92	2.8	446	6	AX127764	Sequence
32	92	2.8	446	6	AX139751	Sequence
33	92	2.8	9741	6	AX127830	Sequence
34	92	2.8	9741	6	AX139817	Sequence
35	92	2.8	9741	6	AX351038	Sequence
36	65	2.0	126295	9	AC004139	Homo sapi
37	61	1.9	206966	9	AC021850	Homo sapi
38	61	1.9	278572	10	AF287263	Mus muscu
39	58	1.8	57662	2	AC107969	Homo sapi
40	58	1.8	81874	9	HS931E15	Human DNA
41	58	1.8	98360	9	HS01247C2	Human DNA
42	58	1.8	145178	2	AC107939	Homo sapi
43	58	1.8	146312	2	AC015494	Homo sapi
44	58	1.8	155982	2	AC015495	Homo sapi
45	58	1.8	157599	2	AC018988	Homo sapi

ALIGNMENTS

RESULT	1	AX351029	3231 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX351029	Sequence 1 from Patent WO0183746.				
DEFINITION	AX351029	Sequence 1 from Patent WO0183746.				
ACCESSION	AX351029	GI:18616385				
VERSION	AX351029.1					
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	1 (sites)					
AUTHORS	Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.					
TITLE	Regulatory nucleic acid sequences of the abcl gene					
JOURNAL	Patent: WO 0183746-A 1 08-NOV-2001;					
FEATURES	Aventis Pharma S.A. (FR)					
source	Location/Qualifiers					
	1..3231					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
BASE COUNT	809 a	773 c	876 g	773 t		
ORIGIN						

Query Match		100.0%;	Score 3231;	DB 6;	Length 3231;		
Best Local Similarity		100.0%;	Pred. No. 0;				
Matches 3231;		Conservative	0;	Mismatches	0;	Indels	0; Gaps
		0;					
Qy	1	acaggcagtggtgagcagtgctgtaatactcagttactcggagggtgaggttgcaatga	60				
Db	1	ACAGGGCATGGTGCAGCGTGCCTGTAATCTCAGTTACTCGGGAGGTGGAGGTTGCAATGA	60				
Qy	61	gccagatcgccacattgcactccagcctgggcaaaaaagggtgaactccatccaatt	120				
Db	61	GCCAGATCGCACCATTTGCACTCCAGCTCGGCAACAAAAGGTGAACACTCCATCTCAATT	120				
Qy	121	aaaaaaaagaatgatttgggtgcactccaattagtagaggaagaagaagagagag	180				
Db	121	AAAAAAAAGATGATTTTGGTGGTGCATCTCAATATAGGTAGGAAGGAGAGAGG	180				
Qy	181	agatggaggtcgaggagatctataactctctaaatcatgctaggaagataaacacct	240				
Db	181	AGATGGAGGTCAGGGAGATCTAATTACTCTTAAATCATGCTAGGAAGATGAACACCT	240				
Qy	241	ttaataaacactctctgttttataacatcattctgccaggagctcaagggtttcaaca	300				
Db	241	TTTAATAACACTCTCTGCTTTTATAACATCATTTCTGCCAAGGAGCTCAAGGTTTCAAGA	300				
Qy	301	aagttcactttcagaaaaccccttgaggagacagaataatacatctctctccatttta	360				
Db	301	AAGTTCATTTTCAGAAAACCCCTTTGAGGAAGACAGAAATATACATCTTCTCTCCATTTTA	360				
Qy	361	aagatgaagaacaggccgggacaaatggctaatgctgtaatccccagcacctttggagg	420				
Db	361	AAGATGAAGAAGACAGGCGGGGCAATGGCTAATGCTGTAAATCCAGCACCTTTGGGAGG	420				
Qy	421	ctaaggccagagatcgcttgactccagagtttgacagcctggatgaataatggcaaa	480				
Db	421	CTAGGCCAGAGGATCGCTTGACTCCAGAGTTTGACAGCAGCCTGGATTAACATGGCAAA	480				
Qy	481	acctgtctctacaaaaataacaaaaattagatgggtggtggtgcacgtgctgtgt	540				
Db	481	ACCTGTCTCTACAAAAAATAACAAAAATTAGATGGGTGTGGTGCATGCACCTGTGTT	540				
Qy	541	ccagctacttgaggagtaagggtgaggtatcgcttgagccaggagtgcaagtctaca	600				
Db	541	CCAGCTACTTTGGGAGGCTTAAGGTGGGAGGATCGCTTGAGCCGAGGAGTCAAGTCTACA	600				
Qy	601	ctgagccatgattggatcactgcactccagcctgggtgagacagacagacccctgtc	660				
Db	601	CTGAGCCATGATTTGGATCACTGCTCAGCTCCAGCCGCGGTGACAGAGCAAGCCCTGTCTCA	660				
Qy	661	aaaaaagaatgaagagaaagaagaagaaggagagagagagagagagagagagag	720				
Db	661	AAAAAAGAAATGAAGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720				
Qy	721	gagggggag	780				
Db	721	GAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	780				
Qy	781	gatgaacag	840				
Db	781	GATGAACAG	840				
Qy	841	ccccaaacccaatttatgacaaaggttattcttgactgaggaagggtccgctct	900				
Db	841	CCCCAAACCCCAATTTATTTGACCAAGTTTATTTCTTTGACTGAGGCAAGGGGTCGCTCT	900				
Qy	901	ctgggcttggttgag	960				
Db	901	CTGGGCTTGGCTTGAAGAGCTCATCTCTGGCCCTTTCTGAGATCCATCCCTTCTTT	960				
Qy	961	ttaattttttgacagggaggtctgtctgtcactcaggttgagtgagtgagtgatgc	1020				
Db	961	TTAATTTTCTTGACACGGAGTCTTGCTCTGTCTCAGTCAAGGCTGAGTGCAGTGCATGATC	1020				

Qy	1021	tcgactcactgtaaacctctgctccgggttcaagcagattctctcctcagctccctcga	1080
Db	1021	TGACTCCTACTTAACCTCTGCTCCCGGTTCAAGCATCTCTCCCTCAAGCTCCCTGA	1080
Qy	1081	gatacagcgccgcacacacatctggctaattttgtatttttagtaaaagactgggtt	1140
Db	1081	GATAACAGCGCCGCCACACATCTGGCTAATTTTGTATTTTATTTAGTAAGACTGGGT	1140
Qy	1141	tcatactgtggccaggggttggtttgcgaactcctgacctgaggtgagctgcccacattggc	1200
Db	1141	TCATCATGTGGCCAGGTTGGTTTGCACCTCTGACCTGAGGTGAGCTGCCACCTTGGC	1200
Qy	1201	ctcccaaatgctggattacagcagcatagccactgcgccagctcaatccatcccttt	1260
Db	1201	CTCCCAAGTGTGGGATTACAGGCATGAGGCACCTGCGCCACCTCAGATCCATCCCCTT	1260
Qy	1261	ctaaaggcaaacagtccttggtgcgaaggggccatgcaccacccagagttatgagtacctgg	1320
Db	1261	CTAAGGGCAACAGTCCATGCTGTCGAAGGGCCATGCCACCCAGAGTTATGAGTACCTGG	1320
Qy	1321	gactccgaatctctgctgggtggcctccacatgcactccagggcctgcttgggcttc	1380
Db	1321	GACTCCGAATTTCTTGGCTGGCTGGCTCCACATGCATCTCCAGGGCCTGCTTGGGCTC	1380
Qy	1381	ttctatgctgtctcctgagttgataagaccactgatgtgactacctgagcttgagcc	1440
Db	1381	TTCTATGGCTCTGCTCCTGAGTGTGATAGACCACTGATGTGAGTACCTGGCTTGAACC	1440
Qy	1441	gtggcctggagatcctgttgactgtagcagggggcttctgagctgaaatgctctga	1500
Db	1441	GTGGCTGGAGATCCTGTTGACTGTAGCATGTAGGAGGGGCTTGTGCAGCTGAATGCTGCA	1500
Qy	1501	tgcaaggtgggtggagttctggaataatgatgagctggagggtgggaagaagagtgagcttg	1560
Db	1501	TGCAGGTGGTGGGAGTTCTTGAATATGATGGAGCTGGAGGTGGGAAGAAGTAGGCTTG	1560
Qy	1561	gggcagctctctcatgcacacctcattctggccaaactcaggtcaaacctgtaagagctct	1620
Db	1561	GGGCAGCTCTCTCATGCCACCTCATCTTGGCCAAACTCAGGTCAACTGTGAAGAGTCT	1620
Qy	1621	aaatgtaatctgacctcgaagggtgctcaagagatctttgtcaagtgaggagacctt	1680
Db	1621	AAATGTGAATCTGCCCTTCAAGGTGGCTTCAAAAGATATCTTTGTCAAGGTAGGACCTT	1680
Qy	1681	gtggcctccagctgcactccagggcctgctggcctcttctacgggtctgctcctgagt	1740
Db	1681	GTGGCCTCCAGCTGCATTCAGGGGCTTGTGGGCTTCTTCTACGGGTCTGTCTGAGT	1740
Qy	1741	cttctataatccttcagggcagattcattatagactcttcacagttttgacctgagttt	1800
Db	1741	CTTCTATGAATCTTCAGGGCAGATTTCATATTAGACTCTTTCACAGTTTGACCTGAGTTT	1800
Qy	1801	tggccagaataagtgacatttagttgtgcttgatgagtgacttaaatattatagaca	1860
Db	1801	TGGCCAGAAATAGGTGACATTTAGTTTGTGGCTGTGATGACTTAAATATTATAGACA	1860
Qy	1861	tgggtgtagcctgcactcctactcttgccttttttttgcctccagtggttttgggt	1920
Db	1861	TGGTGTGTAGGCTTGCAATCTCTCTGCTTTTGTGGCTTGTGGCTTGTGGCTTGTGGGT	1920
Qy	1921	agtttctcctccctacagcgaagaagaagaagaagaagaagaagaagaagaagaaga	1980
Db	1921	AGTTTCTCCTCCCTACAGCCAAAGGCAAAAGAGAGTTGGAGGCTCTGGAGTGGCTACAT	1980
Qy	1981	aattttacagagctgcaattctctgctgacattcacaataatgatacaaaactaaatacaa	2040
Db	1981	AATTTTACAGACTGCAATTTCTCTGGGTGCTTTCACAAATGTATACAACTTAAATACAA	2040
Qy	2041	gtcctgtgtttttatcacagggaggtgataataataataataataataataataataata	2100
Db	2041	GTCCTGTGTTTTATCACAGGGAGGCTGATCAATATATGAATTAAGAAATTAAGGGGCTGGTC	2100

/number=9
 125285..125424
 /gene="ABCA1"
 /number=10
 125754..125870
 /gene="ABCA1"
 /number=11
 130079..130276
 /gene="ABCA1"
 /number=12
 131023..131228
 /gene="ABCA1"
 /number=13
 131749..131925
 /gene="ABCA1"
 /number=14
 133708..133930
 /gene="ABCA1"
 /number=15
 135677..135898
 /gene="ABCA1"
 /number=16
 136959..137163
 /gene="ABCA1"
 /number=17
 138268..138381
 /gene="ABCA1"
 /number=18
 140179..140350
 /gene="ABCA1"
 /number=19
 141340..141471
 /gene="ABCA1"
 /number=20
 142777..142919
 /gene="ABCA1"
 /number=21
 143123..143260
 /gene="ABCA1"
 /number=22
 143963..144183
 /gene="ABCA1"
 /number=23
 145441..145513
 /gene="ABCA1"
 /number=24
 146657..146859
 /gene="ABCA1"
 /number=25
 148527..148575
 /gene="ABCA1"
 /number=26
 148771..148884
 /gene="ABCA1"
 /number=27
 150280..150428
 /gene="ABCA1"
 /number=28
 152078..152202
 /gene="ABCA1"
 /number=29
 153438..153536
 /gene="ABCA1"
 /number=30
 156568..156757
 /gene="ABCA1"
 /number=31

Query Match 54.1%; Score 1749; DB 9; Length 201144;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 qggcatggtggcagggtgctgtgtaactctcaggtgagggtgagggtgcaatgagcc 63

Db	31069	GGGCATGGTGGCAGGTGCCTGTAATCTCAGGTGGAGGTGCAATGAGCC	31128
Qy	64	cagatcgaccattgcactccagcctgggcaacaagtgaactccatctcaataaa	123
Db	31129	CAGATCGACCATTCAGTCCAGCCTGGGCAACAAAGGTGAACCTCCATCTCAATAA	31188
Qy	124	aaaaaagaatgattttggtggtcgacttcaaatagtaggagaagaagagagaga	183
Db	31189	AAAAAAGAATGATTGGTGGTCGACTTCAATAGTAGGAGAAGAGAGAGAGA	31248
Qy	184	tggagggtcaggagatcctaattactctctaaatcatgctaggaagaatacacctttt	243
Db	31249	TGGAGGTCAGGAGATCTAATTAATCTCTTAAATCATGCTAGGAAGATTAACACCTTTT	31308
Qy	244	ataaacactctctgcttttataacatcattctccaaagagctcaaaaggtttcaacaaag	303
Db	31309	ATAAACACTCTCTGCTTTTATAACATCATCTGCCAAGAGCTCAAAAGTTTCAACAAAG	31368
Qy	304	ttcactttcagaaaaacccctttgagaagacagaataatacatcttctctocatttaaag	363
Db	31369	TTCACTTTTCAGAAAAACCCCTTTGAGGAAGACAGAATATACATCTTCTCCATTTTAAAG	31428
Qy	364	atgaagaaacagccgggcaaatggtctaactgctgaatccccagcactttgggaggtcg	423
Db	31429	ATGAAGAAACAGCCGGGCAATGGCTAATGCTGTGTAATCCCCAGCACCTTGGGAGGCTG	31488
Qy	424	agccagagatcgcttgagctccagagtttgagaccagcctgggataacaatggcaaaacc	483
Db	31489	AGCCAGAGGATCGCTTGAGCTCCAGAGTTTGAGACCAGCTGGATTAACATGTCAAAAACC	31548
Qy	484	ctgtctctacaaaaataacaaaaatagatgggtggtggtgagcagcacttggtccccc	543
Db	31549	CTGTCTCTACAAAAAATAACAAAAATTAGATGGGTGTGGTCATGCACCTGTGTCTCC	31608
Qy	544	agctacttggaggctaaagtggagagatcgcttgagccccagggaggtcaagtctcaactg	603
Db	31609	AGCTACTTGGGAGGCTAAGGTGGGAGGATGCTGTGACCCAGGGAGTCAAGTCTACACTG	31668
Qy	604	agccatgattggatcactgcactccagcctcctgggttagacagacagaccctgtctcaaaa	663
Db	31669	AGCCATGATTGGATCACTGCATCCAGCTCCAGCTGGGTAGACAGCAAGACCCTGTCTCAAAA	31728
Qy	664	aaaagaaatgaagagaagaagaagaagagagagagagagagagagagagagagag	723
Db	31729	AAAAGAAATGAAG	31788
Qy	724	ggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag	783
Db	31789	GGGGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAT	31848
Qy	784	gaacagagggcagaagaactttacgttaaatgctcatcatgtgttcaagtgtgacct	843
Db	31849	GAACAGAGGCAGAAAGACTTTACGTAAATGTCTCATCATGTGTGTGTGTGTGTGTGTGT	31908
Qy	844	caaaacccaatttatgaccaggttatctttgactgaggaaggggggttcgctctcct	903
Db	31909	CAAAACCCCAATTTATTTGACCAAGGTATCTTTTACTGAGCAAGGGGGTCCGCTCTCCT	31968
Qy	904	gggcttgggctttagaagctcatctctggcctttctgagatccatccctttctttaa	963
Db	31969	GGGCTTGGGCTTTAGAAAGCTCATCTCTGGCCCTTCTTGAGATCCATCCCTTCTTTTAA	32028
Qy	964	ttttcttgacacggagctctgctctgtcactcaggtggagtgagtgagtgagtgagtg	1023
Db	32029	TTTTTCTTGACACGGAGTCTTGTCTGTCTACTCAGGCTGGAGTGCAGTGGCATGATCTCG	32088
Qy	1024	actcaactgaacctctgctcccggttcaagcattctcctcctcagcctcctcagat	1083
Db	32089	ACTCACTGAACCTCTGCTCTCCCGGTTCAAGCGGATTTCTCTGCTCAGCCCTCTGAGAT	32148
Qy	1084	aacaggcccccgcaccacatctggttaattttgtatttttagtaaaagactgggttca	1143

Db 32149 AACAGGCCGCCACACACTCTGGCTAAATTTTGTATTTTAGTAAGACTGGTTTCA 32208
Qy 1144 tcatgttgccaggttggtttcgaactctgacctgaggtgagctgccacctggcctc 1203
Db 32209 TCATGTTGCCAGGTTGTTTCGAACTCTGACCTGAGGTGAGCTGCCACTTGGCCTC 32268
Qy 1204 ccaagtgtcgtggattacagagcatgagccactgccccagctcagatccatccctttcta 1263
Db 32269 CCAAGAGTCTGGGATTACAGGACATGAGCACTGGCCCGCAGTCAGATCCATCCCTTTCTA 32328
Qy 1264 agggcaaacatccatggtgcacaggggccatgccccagctcagatccatccctttcta 1323
Db 32329 AGGGCAACAGTCCATGTTGTCAGAGGCGCCATGCGACCCAGCTTATGAGTACCTGGGAC 32388
Qy 1324 tccagaattctctgctggtgacctccacatgcaactccagggcctctggcctcttc 1383
Db 32389 TCCAGATTCTTCCCTGGTGGCTCCACATGCACTTCCAGGGCCTCTTGGGCTCTTC 32448
Qy 1384 tatgctgtctcgtgagttgtatagaaacactgagtgtgactgctgggttgagccgtg 1443
Db 32449 TATGCTGTCTGCTGAGTGTGTATAGAACCACTGATGTGAGTACCTGGCTTGAGCCGTG 32508
Qy 1444 gctcggagatccttaccatgtagctgagggggtgtgagctgaatgtctgcatgc 1503
Db 32509 GCCTGGAGATCCTCTTGACTGTAGCATGGAGGGGCTTGTGCAGCTGAATGCTGCTGATGC 32568
Qy 1504 aggtggtgaggtctggaatatgtagagctgaggtggaagagagtagcttgggg 1563
Db 32569 AGTGTGTGGAGTCTGGAATATGATGAGCTGGAGTGGAGAGTGGCTTGGGG 32628
Qy 1564 cagctctctcagcactcattctggccaaactcagggtcaactgtgaagagctctaaa 1623
Db 32629 CAGCTCTCTCATGCACTCATTCTGCCAAACTCAGGTCAACTGTGAAGTCTAAA 32688
Qy 1624 tgtgaatctgcccctcaggtggtacaaagggtatcttctcaggtgagagacctgtg 1683
Db 32689 TGTGAATCTGCCCTCAAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTG 32748
Qy 1684 gctccacgtgacctccagggcctgctggcctctctcagggctgctcagctctt 1743
Db 32749 GCCTCCAGGTGCACTTCCAGGGCTGTTGGGCTCTTACGGGTCTGCTGAGTCTT 32808
Qy 1744 ctatgaatc 1752
Db 32809 CTATGAATC 32817

RESULT 4
AF275948 149034 bp DNA linear PRI 17-JUL-2000
LOCUS Homo sapiens ABCAL (ABCA1) gene, complete cds.
DEFINITION AF275948
ACCESSION AF275948.1 GI:9247085
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 149034)
Santamarina-Fojo, S., Peterson, K., Knapper, C., Qiu, Y., Freeman, L.,
Cheng, J.F., Osorio, J., Remaley, A., Yang, X.P., Haudenschild, C.,
Prades, C., Chimin, G., Blackmon, E., Francois, T., Duverger, N.,
Rubin, E.M., Rosier, M., Deneffe, P., Fredrickson, D.S. and Brewer, H.B.
Jr.
Complete genomic sequence of the human ABCAL gene: analysis of the
human and mouse Atp-binding cassette A promoter
Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
20345099
2 (bases 1 to 149034)
Santamarina-Fojo, S., Peterson, K.M., Knapper, C.L., Freeman, L.A.,
Remaley, A.T., Yang, X.-P., Haudenschild, C.C., Blackmon, E.E.,
Francois, T.D. and Brewer, H.B. Jr.
Direct Submission

JOURNAL Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA
FEATURES
Source Location/Qualifiers
1. 149034
/organism="Homo sapiens"
/db_xref="taxon:9606"
Join(1454, 1674, 25831, 25989, 40385, 40478, 45012, 45153,
46423, 46541, 67729, 67850, 70831, 71007, 83960, 84052,
89010, 89250, 91962, 92101, 92433, 92549, 96758, 96955, 97702,
97702, 97907, 98428, 98604, 100391, 100613, 102360, 102581,
103642, 103846, 104951, 105064, 106862, 107033,
108023, 108154, 109460, 109602, 109806, 109943,
110646, 110866, 112124, 112196, 113183, 113385,
115053, 115101, 115297, 115410, 116806, 116954,
118604, 118728, 119964, 120062, 123094, 123283,
124804, 124898, 126208, 126240, 127363, 127468,
128943, 129017, 129539, 129708, 130939, 131116,
133114, 133229, 133341, 133485, 134526, 134649,
135736, 135865, 136130, 136250, 137038, 137100,
138008, 138114, 140469, 140610, 140982, 141116,
142060, 142163, 142646, 142738, 143397, 143640,
144581, 148034)
/gene="ABCA1"
/product="ABCA1"
Join(1454, 1674, 25831, 25923)
/gene="ABCA1"
1454, 148034
/gene="ABCA1"
2406, 2678
/rpt_family="Alu"
4240, 4509
/rpt_family="Alu"
4998, 5270
/rpt_family="Alu"
5563, 5839
/rpt_family="Alu"
6615, 6877
/rpt_family="Alu"
8800, 9046
/rpt_family="Alu"
9998, 10279
/rpt_family="Alu"
11859, 12133
/rpt_family="Alu"
12810, 12902
/rpt_family="Alu"
15220, 15399
/rpt_family="Alu"
15508, 15787
/rpt_family="Alu"
18600, 18860
/rpt_family="Alu"
20950, 21206
/rpt_family="Alu"
25038, 25314
/rpt_family="Alu"
Join(25924, 25989, 40385, 40478, 45012, 45153, 46423, 46541,
67729, 67850, 70831, 71007, 83960, 84052, 89010, 89250,
91962, 92101, 92433, 92549, 96758, 96955, 97702, 97907,
98428, 98604, 100391, 100613, 102360, 102581, 103642, 103846,
104951, 105064, 106862, 107033, 108023, 108154,
109460, 109602, 109806, 109943, 110646, 110866,
112124, 112196, 113183, 113385, 115053, 115101,
115297, 115410, 116806, 116954, 118604, 118728,
119964, 120062, 123094, 123283, 124804, 124898,
126208, 126240, 127363, 127468, 128943, 129017,
129539, 129708, 130939, 131116, 133114, 133229,
133485, 134526, 134649, 135736, 135865,
136130, 136250, 137038, 137100, 138008, 138114,
140469, 140610, 140982, 141116, 142060, 142163,
142646, 142738, 143397, 143640, 144581, 148034)
/gene="ABCA1"


```

/codon_start=1
/product="ABCA1"
/protein_id="AAF86276.1"
/db_xref="GI:9247086"

/translation="MACWPQLRLWLKNTLRRRTQCOLLEAVAMPFLIFILLISVRL
SYPPYEQHECHFPKAMPAGTLPWQGIICNANPCFRYPPTGEAFVGVGNFKNSIV
ARLFSDARLLYSQKDTSMKMRKRLVRLTQIIRKSSNLKLODFLVDNETFSGFLYH
NLSLPSKTVDMKLRADVLSKMFVQLQYOLHSLCNGSKSEMIQLGDOEVSFGLCP
REKLAARVLRNSNDILKPLILRLNSTSPFPEKELAPATKTLHSLGLTQAEQFSMR
SHSDMRQVEMLTNNSSSSTQIYQAVSRIVCGHPGEGGLKIKSLWYEDNNYKALF
GNGTEAEFIYDNLNKNLESSPLIKWALKPLLVGLKILITPDPAT
ROYMAEVNKTQELAVFHLEPMWELSPKWTWENSOEDMLVRLDLSRDMFWHE
QQDLGDTAQDIAVFLAKHPEDVQSSNGSVTWAEFNETNOAIRISREMECVNL
KLEPIATFVILINKSMELLDERKFWAGIVFTGIPGSELPHVYKIKRMDINVERT
NKIKDGVNDPGRFEDRMVGMGRAYLODVVEQAIIRVLTGTEKTKYVYMOOMP
PCYVDIIFLWMSRSMPLMLANIYSVAVILKIVYEKEARLKMTRIMGLDLSILW
FSWFTSSILPLVSLAGLIVILKGNLLPYSDPSVVFVFLSVFAVVLQCFLLSTLF
SRANLAAACGGIYFTLPLVYLCVADQDYFTLIFAXLLSPVAFGFGCEYFALFE
EGGIGQVNDLFPESDEEDFNLTTSVMMLFDTLFGVMTWYIEAVPPGYGIPRW
YPCTKSYWFGFESDESHGNSQKRISEICMEBEPHKLGVSTQNLVYVRDGMKV
AVDGLALNFEQITSLFNGHNGAGKTMTMSILTLGFLPPTSGTAVILGKDISENSTIR
ONLGVCPQHNVLFDMLTVEEHWFYARLKGISEKHVKAEMEQMALDVLPSKLSKT
SOLSGMQRKLSVALAFVGGSKVILDEPTAGVDYPSRGTWELLKXQRQRTILST
HMDLADVLGDRALISHLKGLCCVGSGLFLANQUGTYLTLVKKDVSESSCRNS
STVSLKXEDSVQSSDAGLSGDSHESDTLTDVSAISNLKRVSEARLVEDIGHEL
TYVLPYAAKEGAFYELPHEIDRLSDIGISSYGETTLEEFILKVAEESGVDAETS
DCTLPAARNRRAFGQKQLRPFEDDAADNDSIDIPESRETDLGMDGKGYQVK
GKLTQOQFVALLKRLIARRSRKGFPAQIVLPAVFCIALVFSLIVPPPKGYPSLE
LQPMWYNOQYTFVNDAPEDGTLELLNALIKDGFCTRCMEGNPIPTPCQAGEEW
TTPVPTQITMDLFGNGMTQNPSPACOCSSDKIKMLPVCPPGAGGLPPQKQNTA
DILQDQTKRNLSDLYVTVQVLIASLKNKIWNENFRYGGFSLGVSNTQALPPSOEYN
DAXKQMKHLKLKADSSADRLNLSLGRMTGLDTRNNKVMFNKNGHAISSFLNVIN
NAILRANLQKNSHYGHTAFNHPNLTKOOLSEVAXMTTSDVVLVSIIVFAMSFV
PASVFLVIOERVSRAKHLORTISGVKPVYIWLNSFVDMCNVVPATVLIIVIFICFOQ
KSYVSTNLPVALLLLLSWISITPLMYPASFPVKIPSTAYVVLTSVNLFGINGSVA
TVLELFTDKNNLNDILKSVLIFPHFCGLRGLIDMVKNQAMADALEREGENRFS
PLSBDLVLFUNNAVEGVVFLITVLIQYRFFIRPRVNAKLSPLNDEDEDVRRERO
RLDGGQNDILEIKELKVIYRRKRKPAVDICVIGPEGCGFLGLVNGAGKSSFTFM
LGDGTVTVRGDAFLNKNLSILNIHEVHQNMCYCPOFDAITELLGTHREHVEFALLRGV
PKPEKYGEMAIRKGLVYKGEKAGNYSGNKRKLSTAMALIGGPPVPLDDEPTTG
MDKARRELWNCALSVKREGSVLTSMSRECEALCTRMALMVNGRFRCLGSHOHLK
NRFQGYITVIRAGSNFPLKVPQDFGLAFPGSVYKKEKHNMLQYQLPSSLSLARI
FSLQSKRKLKHEDYSVSQTLQDQVFNFAKQDSDDHLKLSLHKNQTVVDVAULT
SFLODEKVESYV"

repeat_region 29916..30180 /rpt_family="Alu"
repeat_region 31050..31314 /rpt_family="Alu"
repeat_region 31692..31973 /rpt_family="Alu"
repeat_region 36167..36440 /rpt_family="Alu"
repeat_region 37617..37892 /rpt_family="Alu"
repeat_region 37999..38166 /rpt_family="Alu"
repeat_region 38172..38269 /rpt_family="Alu"
repeat_region 38995..39271 /rpt_family="Alu"
STS 41547..41565 /gene="ABCA1"
repeat_region 41593..41835 /note="A009X28"
STS 41657..41679 /rpt_family="Alu"
repeat_region 42320..42584 /note="A009X28"
repeat_region 44349..44626 /rpt_family="Alu"
repeat_region 48473..48669 /rpt_family="Alu"

Query Match 46.0%; Score 1487; DB 9; Length 149034;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1787; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1445 cctggagatcctgttgacttagcatgagggggcctgtgagcagtgatgctcatgca 1504
DB 1 CCTGGAGATCTGTTGACTGTAGCATGGAGGGGCTTGTGCAGCTGAATGTCTGCATGCA 60
QY 1505 ggtggtggagttctgggaatgatgagagctggaggtgggagagagagcttgaggc 1564
DB 61 CGTGTGGGAGTCTCGGAATATGATGAGCTGGAGCTGGAGGGAAGAGACTAGCTTGGGC 120
QY 1565 agctctctatgccacacctcctctctggcccaaaactcgaagctggaagagctctaaat 1624
DB 121 AGCTCTCTCATGCCACCTCATCTTGGCCAAAACCTCAGGTCAAACTGTGAAGAGCTAAAT 180
QY 1625 gtgaatctgcccttcaagtgagtgctacaagaagtattcttgcaggtacagtaggagaccttg 1684
DB 181 GTGAATCTGCCCTTCAAGGTGGCTACAAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGG 240
QY 1685 cctccacgtgcactccaggcctgcttggcctctctacgggtctctcctgagctcttc 1744
DB 241 CCTCCACGTGCACCTTCAGGGCCCTGCTTGGCCCTCTTCTACGGGTCTGCTCCTGAGCTTC 300
QY 1745 tatgaa----tcttcaggggcagattcattattagactcttcacagtttgacctgagttt 1800
DB 301 TATGAATCTGCTTCTCAGGGCAGATTTCATATTAGACTCTTACACAGTTTGACCTGAGTTT 360
QY 1801 tggcagagaataagtgacatttagtttggcttgatgagtaggacttaataattatagaca 1860
DB 361 TGGCCAGAAATAAGGTGACATTAGTTTGTGGCTTGTATGGATGACTTAAATATTATTAGACA 420
QY 1861 tgggtgtaggcctgcattcctactcttgcctcttttttttgcctccctcagtgcttttgggt 1920
DB 421 TGGTGTGTAGGCTGCATCTCTACTCTTGGCTTTTGTTCCTCCAGCTGTTTGGGT 480
QY 1921 agttttgtccctacagccaaagagagagagagagagagagagagagagagagagagagag 1980
DB 481 AGTTTGTCTCCCTACAGCCAAAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 1981 aattttacagcactgaattctctggtgctgacattcacaaatgatacaaaacttaatacaaa 2040
DB 1981 aattttacagcactgaattctctggtgctgacattcacaaatgatacaaaacttaatacaaa 2040

```



```

misc_feature      84249..84273
/note="Sequence from overlapping clone RP11-122P10
(Q026643). Assembly confirmed by restriction digest."
92050..92163
/misc_feature     /note="Sequence from reads from a short insert library
derived from a single pUC clone. Restriction digest data
confirm the assembly."
92411..92557
/misc_feature     /note="Sequence from reads from a short insert library
derived from a single pUC clone. Restriction digest data
confirm the assembly."
BASE COUNT       27673 a 21138 c 20380 g 27526 t
ORIGIN
Query Match              27.6%; Score 891; DB 9; Length 96717;
Best Local Similarity    99.7%; Pred. No. 0;
Matches 1041; Conserved   0; Mismatches 3; Indels 0; Gaps 0;

Qy 709 gagggagaggaggggaggaagaagaagaagaagaagaaaaaaatgataaaa 768
Db 19434 GAGGSGAGGGAGGGGGAGGAAGAAGAAAGAAAAGAAAAAATGATAAAA 19375

Qy 769 agaaaaaaaaaacagtgaacacagagcgcaagaagactttacctcatctgtgt 828
Db 19374 AGAAAAAAAAACAGATGAACAAGCAGCGCAAGAACACTTTACGTAAATTGCTCATCATGTGGT 19315

Qy 829 tgtcaagtgtaaccaccaaaacccaatttatgaccaaggttattcttgactgaggaag 888
Db 19314 TGTCNAAGTTTTGACCCCAAACCCCATAATTATTATGCACCAGGTTATTCTTTTGACTGAGGCAAG 19255

Qy 889 gggtgcgcgtcctcctgggcccttggcccttagaaagctcatctcggcccttcctaagatccc 948
Db 19254 GGgGTCCGCTCTCCTGGGCCCTTGGGCTTTAGAAAGCTCATCTCGGCCCTTTCTGAGATCC 19195

Qy 949 atccctttctttttatttttttttttacacggaggtcttgtctgtcacctcaggctggagtgc 1008
Db 19194 ATCCCTTTCTTTTATTATTTTCTTTGACACGGAGTCTTGCTGTCTGTCTCACTCAGGCTGGAGTGC 19135

Qy 1009 agtggcatgatctcgactcactgtaaacctctgcctccccgggttaacggattctctgc 1068
Db 19134 AGTGcATGATCTCGACTCATACTGTAACTCTGCCCCGGGTTCAAGCGATTCCTCTGCC 19075

Qy 1069 tcagcctcctgagataacagcgccgcgcacacacatctggcataattttgtatttttagt 1128
Db 19074 TCAGCCTCCTGAGATAACAGCGCTCGGCCACCACATCTGGCTAATTTTGTATTTTAGT 19015

Qy 1129 aaagactgggtttcatcatatgttgccagggttggttttcogaaacctcctgaoctgagtgagct 1188
Db 19014 AAAGACTGGGTTTTCATCATGTTGGCCAGGTTGGTTTTCGAACTCCTGTGACTCGAGTGAGCT 18955

Qy 1189 gccacacttggcctcccaaagtctcgaggtacagggatgaagccactgcgcccagctcag 1248
Db 18954 GCCACCTTGGCCTCCCAAAAGTGTGGGATTACAGGCATGAGCCACATGCGCCCGAGCTCAG 18895

Qy 1249 atccatoccttttaaaggcacaacagtcctcattggtgc aaaaggccatgccaccacagagtt 1308
Db 18894 ATCCATCCCTTTCTAAGGGCAACAGTCCATGGTGCAAAAGGGGCCATGCCACCAGAGTT 18835

Qy 1309 ataagtaacctgggaactccagaattccttgctgggtggcctccacatgcaattccagggcc 1368
Db 18834 ATGAGTAACCTGGGACTCCAGAAATTCCTTGGCTGGTGGCCCTCCACATGCACTTCACGGGCC 18775

Qy 1369 tgcttgggacctcttctatcgctcgtctgactgagtgattgatatagaaccactgattgagtaacc 1428
Db 18774 TGCTTGGGCCTCTCTCATGGGTCTGTCTCTGAGTGTGATAGAACCCTGATGTGAGTACC 18715

Qy 1429 tgggcttyagccgttggcctggagatcctgttgactgtagcatgagggggttggtgagc 1488
Db 18714 TGGGCTTGAGCCGTGGGCTTGAGATCTGTTGACTGTAGCATGGAGSGGGCTTGTCGACG 18655

Qy 1489 tgaatgtctcagtcaggttggtgggagttctggaatatgatgagcttgagggtgggaaga 1548

```

```

Db 27031 CATCATGTTGCCAGGTTGGTTTCGACTCTCTGACCTGAGTGAGCTGCCACCTTGGCC 27090
Qy 1202 tcccaagtgctgggattacagggcatgagccactgcgcccagctccagatccatcccttc 1261
Db 27091 TCCCAAAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGCTCAGATCCATCCCTTTC 27150
Qy 1262 taagggaacacagtcctatggtgcaaaagggcccatgcccacccagagttatagttacctggg 1321
Db 27151 TAAGGGCAACAGTCCATGGTGTGAAGAGGGGCCATGCGCCACGAGTATATAGTACCTGGG 27210
Qy 1322 actccagaattccttgctggtggtccacatgcatctccaggcctgcttggtggcctct 1381
Db 27211 ACTCCAGAAATTCCTTGCTGTGGCTCCACATGTCACCTTCCAGGGCTGCTTGGGCTCT 27270
Qy 1382 tctatgctgctgctgagtggtgatagaaccactgagtgtgagttacctgggtgagcgg 1441
Db 27271 TCTATGGGTCTGTCTGCTGAGTGTGATAGAACCACCTGATGTGAGTACCTGGCTTGAGCCG 27330
Qy 1442 tggcctgagatcctgttgactgtagcatgagggggtgctgtgacgtgaatctgcat 1501
Db 27331 TGGCCTGGAGATCTGTTGACTGTAGCATGAGGGGGCTTGTGCAGCTGAATGCTGTGYAT 27390
Qy 1502 gcaagtggtgggagttctggaatatgtaggagctggaggtgggaagagaagtaggtctgg 1561
Db 27391 GCAGGTGGTGGGAGTCTTGGAAATATGATGAGCTGGAGGTGGAGAGAAAGTAGGCTGG 27450
Qy 1562 ggcagctctctatgcccactcttctggccaaaactcaggtcaaacctgtaagagctcta 1621
Db 27451 GGCAGCTCTCTCAFGCCACCTCAFTCTGGCCAAACTCAGGTCAAACTGTGAAGAGTCTA 27510
Qy 1622 aatgtaactgccttcaagggtggctacaaaggtatctttgtcaaggtagagaccttg 1681
Db 27511 AATGTGAATCTGCCCTTCAAGGTGGCTACAAAGGTATCTTTGTCAAGTAGGAGACCTTG 27570
Qy 1682 tggcctccagtgcaactccagggcctgctgggctcttctacaggtgtgctcctgagtc 1741
Db 27571 TGGCCTCCAGTGCACCTTCCAGGGCCTGCTTGGGCTCTTCTACGGGTCTGTCTCCTGAGTC 27630
Qy 1742 tcttatg 1748
Db 27631 TTCTATG 27637

RESULT 7
AC012230
LOCUS Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
DEFINITION AC012230 175064 bp DNA linear HTG 22-APR-2000
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175064)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-1M10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175064)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Balwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,K., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

```

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-Oct-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2510

Center clone name: 1_M.10

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 117571 bases at least Q40

Consensus quality: 145749 bases at least Q30

Consensus quality: 160940 bases at least Q20

Insert size: 185000; agarose-fp

Insert size: 171264; sum-of-contigs

Quality coverage: 2.9 in Q20 bases; agarose-fp

Quality coverage: 3.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1003: contig of 1003 bp in length

* 1004 1103: gap of 100 bp

* 1104 2634: contig of 1531 bp in length

* 2635 2734: gap of 100 bp

* 2735 4415: contig of 1681 bp in length

* 4416 4515: gap of 100 bp

* 4516 5785: contig of 1270 bp in length

* 5786 5885: gap of 100 bp

* 5886 7879: contig of 1994 bp in length

* 7880 9686: contig of 1707 bp in length

* 9687 9786: gap of 100 bp

* 9787 12253: contig of 2467 bp in length

* 12254 12353: gap of 100 bp

* 12354 15228: contig of 2875 bp in length

* 15229 15328: gap of 100 bp

* 15329 17200: contig of 1872 bp in length

* 17201 17300: gap of 100 bp

* 17301 20131: contig of 2831 bp in length

* 20132 20231: gap of 100 bp

* 20232 22587: contig of 2356 bp in length

* 22588 22687: gap of 100 bp

* 22688 25707: contig of 3020 bp in length

* 25708 25807: gap of 100 bp

* 25808 28184: contig of 2377 bp in length

* 28185 28284: gap of 100 bp

* 28285 31338: contig of 3054 bp in length

* 31339 31438: gap of 100 bp

* 31439 34299: contig of 2861 bp in length

* 34300 34399: gap of 100 bp

* 34400 38318: contig of 3919 bp in length

* 38319 38418: gap of 100 bp

* 38419 42835: contig of 4417 bp in length

* 42836 42935: gap of 100 bp

* 42936 45448: contig of 2513 bp in length

* 45449 45548: gap of 100 bp

*

```

RES      Location/Qualifiers
source   1..175064
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="RP11-1M10"
          /clone_lib="RPC1-11 Human Male BAC"
misc_feature 1..1003
            /note="assembly_fragment"
misc_feature 1104..2634
            /note="assembly_fragment"
misc_feature 2735..4415
            /note="assembly_fragment"
misc_feature 4516..5785
            /note="assembly_fragment"
misc_feature 5886..7879
            /note="assembly_fragment"
misc_feature 7980..9686
            /note="assembly_fragment"
misc_feature 9787..12253
            /note="assembly_fragment"
misc_feature 12354..15228
            /note="assembly_fragment"
misc_feature 15329..17200
            /note="assembly_fragment"
misc_feature 17301..20131
            /note="assembly_fragment"
misc_feature 20232..22587
            /note="assembly_fragment"
misc_feature 22688..25707
            /note="assembly_fragment"
misc_feature 25808..28184
            /note="assembly_fragment"

```



```

QY 3030 ttccccgggttttttcccccttccggaagcctgtcaagggttaggagaaagag 3089
|||||
Db 1063 TTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTGTCAAGGGTAGGAGAAAGAG 1122
|||||
QY 3090 acgcaaacacaaaagtggaaaacagtagtagaggtctctccagtgc 3134
|||||
Db 1123 ACGCAAAACAAAAGTGGAAACAGGTAAGAGGCTCTCCAGTGAC 1167
|||||

RESULT 9
AF258623S1 1167 bp DNA linear PRI 23-JUN-2000
LOCUS
DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,
promoter and exon 1.
ACCESSION AF258623
VERSION AF258623.2 GI:8677405
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE
AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
TITLE Direct Submission
JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
REFERENCE
AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
REMARK
COMMENT Sequence update by submitter
FEATURES
SOURCE
1. .1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/size="224..844"
/gene="ABCA1"
/exon 845..1147
/promoter
BASE COUNT 278 a 313 c 328 g 248 t
ORIGIN

Query Match 18.2%; Score 589; DB 9; Length 1167;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 999; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 2130 ttgtttgtttttttttttgtttgttgcctctctctcctcaatttatgaagaagca 2189
|||||
Db 168 TTGTGTTTGTCTTTTGTGTTTGTGCGCTCTCTCTCAATTATGAAGAAGCA 227
|||||

QY 2190 gtaagatgtctctcgtggtctctgagggacctggggagctcaggctggggaatctcaa 2249
|||||
Db 228 GTAAGATGTTCTCTCTGGGTCCTCTGAGGAGCTGGGAGCTCAGGCTGGGAATCTCCAA 287
|||||
QY 2250 ggcagtaggtgcctatacaaaatacaagctcaggtttgttggtgggggaaacaaagcgc 2309
|||||

```

```

Db 288 GGCACTAGTGCCTATCAAAAATCAAAAGTCCAGGTGTGTGGGGGAAAAACAAAAGCAGC 347
QY 2310 ccattaccagagactgtccgccttccctccacccagcctagccttttgaaggaagac 2369
|||||
Db 348 CCATTACCAGAGAGACTGTCCGCTTCCCTCACCCAGCCTAGCCCTTTGAAGAAGAAC 407
|||||
QY 2370 aaaagacaagacaaaatgattggcgtcctgagggagattcagccttagagctctctctccc 2429
|||||
Db 408 AAAAGACAAGACAAAATGATTGGTCTCTGAGGGAGATTACGCTAGAGCTCTCTCTCCC 467
|||||
QY 2430 ccaatccctccctcgcgtcaggaagaaactaacaaggagaaaaaaatgcggaaagcagga 2489
|||||
Db 468 CCAATCCCTCCCTCCGCTGAGGAACCTAACAAAGGAAAAAATTCGGGAAAGAGGA 527
|||||
QY 2490 tttagaagaagcaaatccactggtcccttgctgctgggaaacgtgagactagagagctct 2549
|||||
Db 528 TTTAGAGGAAGCAAAATTCACCTGGCTGGCTGGCGGGAACGTGGACTAGAGAGTCT 587
|||||
QY 2550 gcggcgagccccgagccccagcgtctcccgcgctcttaggcggcgggccccggcgggg 2609
|||||
Db 588 GCGGCGCAGCCGCGAGCCAGCGCTTCCGCGCGCTTAGCGCGCGCGCGCGCGCGGG 647
|||||
QY 2610 gaagggaagcagagaccgagaccctaagacacctgtacctccacccccacccccaccc 2669
|||||
Db 648 GAAGGGGACGACAGACCGGGGACCTTAAGACACCTGCTGTACCTCTCCAC-----CCCAACC 702
|||||
QY 2670 ccaccacacctcccccaactccctagatgtgctggtggcggtgaaacgtgcgcgttttaa 2729
|||||
Db 703 CCACCCACCTCCCGCCCACTCCCTAGATGTGCTGGGCGGCTGAACGTCGCCCGTTTAA 762
|||||
QY 2730 gggcgggccccggcctccacgctgttctgtgagtagtaactatacaaaagagacc 2789
|||||
Db 763 GGGGCGGCGCGCGCTCCACGTGCTTCTGTGAGTGACTGAACCTAGATAAACAGAGAGCC 822
|||||
QY 2790 ggaagggggcgaggagagagacacagcgtttaccgcatagtaacctctgcctcgc 2849
|||||
Db 823 GGGAAAGGGGGGGGAGGAGGAGGAGACAGAGCTTTGACCGTAGTAACTCTCGCTCG 882
|||||
QY 2850 gtgcagcgcaatctataaaagaaactagtcgccgcaaaaccccgtaattgcggagcgaga 2909
|||||
Db 883 GTGCAGCGCGAATCTATAAAGGAAGTACTGTCTCGCAAAACCCCGTAATTGCGAGCGAGA 942
|||||
QY 2910 gtgagtgggggcgaggaccgagagccgagaccctctctcccggtgcggcgaggg 2969
|||||
Db 943 GTGAGTGGGGCGGACCGCGAGAGCGAGCCCTTCTCTCCGCGGTGCGCGCAGGG 1002
|||||
QY 2970 caggcggggagctcccgcgacacacagcggcttctcaggcgcttgcctctgttt 3029
|||||
Db 1003 CAGGCGGGGAGCTCCCGCGCACCAACAGAGCGGTTCTCAGGGCGCTTGTCTCTCTTT 1062
|||||
QY 3030 ttccccgggttctgttttctccctctccggaaggcttgcgaagggttaggagaaagag 3089
|||||
Db 1063 TTTCCCGGTTCTGTTTCTCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAG 1122
|||||
QY 3090 acgcaaacacaaaagtggaaaacaggttaagaggtctctccagtgc 3134
|||||
Db 1123 ACGCAAAACAAAAGTGGAAACAGGTAAGAGGCTCTCCAGTGAC 1167
|||||

RESULT 10
AX060715
LOCUS
DEFINITION Sequence 3 from Patent WO0078972.
ACCESSION AX060715
VERSION AX060715.1 GI:12406104
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lawn, R.M., Wade, D. and Garvin, M.
TITLE Regulation with binding cassette transporter protein abcl

```


Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lechokzy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheders, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6705871.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L2512
 Center clone name: L_N10

* NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 871: contig of 871 bp in length
 * 872 971: gap of 100 bp
 * 972 1834: contig of 863 bp in length
 * 1835 1934: gap of 100 bp
 * 1935 2804: contig of 870 bp in length
 * 2805 2904: gap of 100 bp
 * 2905 3745: contig of 841 bp in length
 * 3746 3845: gap of 100 bp
 * 3846 4696: contig of 851 bp in length
 * 4697 4796: gap of 100 bp
 * 4797 5640: contig of 844 bp in length
 * 5641 5740: gap of 100 bp
 * 5741 6540: contig of 800 bp in length
 * 6541 6640: gap of 100 bp
 * 6641 7509: contig of 869 bp in length
 * 7510 7609: gap of 100 bp
 * 7610 8479: contig of 870 bp in length
 * 8480 8579: gap of 100 bp
 * 8580 9430: contig of 851 bp in length
 * 9431 9530: gap of 100 bp
 * 9531 10376: contig of 846 bp in length
 * 10377 10476: gap of 100 bp
 * 10477 11322: contig of 846 bp in length
 * 11323 11422: gap of 100 bp
 * 11423 12302: contig of 880 bp in length
 * 12303 12402: gap of 100 bp
 * 12403 13280: contig of 878 bp in length
 * 13281 13380: gap of 100 bp
 * 13381 14241: contig of 861 bp in length
 * 14242 14341: gap of 100 bp
 * 14342 15196: contig of 855 bp in length
 * 15197 15296: gap of 100 bp
 * 15297 16123: contig of 827 bp in length
 * 16124 16223: gap of 100 bp
 * 16224 17072: contig of 849 bp in length

17073 17172: gap of 100 bp
 * 17173 18041: contig of 869 bp in length
 * 18042 18141: gap of 100 bp
 * 18142 19009: contig of 868 bp in length
 * 19010 19109: gap of 100 bp
 * 19110 19966: contig of 857 bp in length
 * 19967 20066: gap of 100 bp
 * 20067 20921: contig of 855 bp in length
 * 20922 21021: gap of 100 bp
 * 21022 21865: contig of 844 bp in length
 * 21866 21965: gap of 100 bp
 * 21966 22832: contig of 867 bp in length
 * 22833 22932: gap of 100 bp
 * 22933 23780: contig of 848 bp in length
 * 23781 23880: gap of 100 bp
 * 23881 24733: contig of 853 bp in length
 * 24734 24833: gap of 100 bp
 * 24834 25670: contig of 837 bp in length
 * 25671 25770: gap of 100 bp
 * 25771 26621: contig of 851 bp in length
 * 26622 26721: gap of 100 bp
 * 26722 27576: contig of 855 bp in length
 * 27577 27676: gap of 100 bp
 * 27677 28532: contig of 856 bp in length
 * 28533 28632: gap of 100 bp
 * 28633 29492: contig of 860 bp in length
 * 29493 29592: gap of 100 bp
 * 29593 30455: contig of 863 bp in length
 * 30456 30555: gap of 100 bp
 * 30556 31410: contig of 855 bp in length
 * 31411 31510: gap of 100 bp
 * 31511 32368: contig of 858 bp in length
 * 32369 32468: gap of 100 bp
 * 32469 33312: contig of 844 bp in length
 * 33313 33412: gap of 100 bp
 * 33413 34268: contig of 856 bp in length
 * 34269 34368: gap of 100 bp
 * 34369 35204: contig of 836 bp in length
 * 35205 35304: gap of 100 bp
 * 35305 36156: contig of 852 bp in length
 * 36157 36256: gap of 100 bp
 * 36257 37128: contig of 872 bp in length
 * 37129 37228: gap of 100 bp
 * 37229 38083: contig of 855 bp in length
 * 38084 38183: gap of 100 bp
 * 38184 39031: contig of 848 bp in length
 * 39032 39131: gap of 100 bp
 * 39132 40006: contig of 875 bp in length
 * 40007 40106: gap of 100 bp
 * 40107 40967: contig of 861 bp in length
 * 40968 41067: gap of 100 bp
 * 41068 41913: contig of 846 bp in length
 * 41914 42013: gap of 100 bp
 * 42014 42824: contig of 811 bp in length
 * 42825 42924: gap of 100 bp
 * 42925 43776: contig of 852 bp in length
 * 43777 43876: gap of 100 bp
 * 43877 44752: contig of 876 bp in length
 * 44753 44852: gap of 100 bp
 * 44853 45724: contig of 872 bp in length
 * 45725 45824: gap of 100 bp
 * 45825 46643: contig of 819 bp in length
 * 46644 46743: gap of 100 bp
 * 46744 47599: contig of 856 bp in length
 * 47600 47699: gap of 100 bp
 * 47700 48551: contig of 852 bp in length
 * 48552 48651: gap of 100 bp
 * 48652 49485: contig of 834 bp in length
 * 49486 49585: gap of 100 bp
 * 49586 50440: contig of 855 bp in length
 * 50441 50540: gap of 100 bp
 * 50541 51404: contig of 864 bp in length
 * 51405 51504: gap of 100 bp

ACCESSION
VERSION
KEYWORDS
SOURCE

AC012230
AC012230.3 GI:7637254
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.

ORGANISM

LISTENING

REFERENCE

REFERENCES AND AUTHORS

AUTHORS	TITLE
	Abstracts
	107-116
	117-126
	127-136
	137-146
	147-156
	157-166
	167-176
	177-186
	187-196
	197-206
	207-216
	217-226
	227-236
	237-246
	247-256
	257-266
	267-276
	277-286
	287-296
	297-306
	307-316
	317-326
	327-336
	337-346
	347-356
	357-366
	367-376
	377-386
	387-396
	397-406
	407-416
	417-426
	427-436
	437-446
	447-456
	457-466
	467-476
	477-486
	487-496
	497-506
	507-516
	517-526
	527-536
	537-546
	547-556
	557-566
	567-576
	577-586
	587-596
	597-606
	607-616
	617-626
	627-636
	637-646
	647-656
	657-666
	667-676
	677-686
	687-696
	697-706
	707-716
	717-726
	727-736
	737-746
	747-756
	757-766
	767-776
	777-786
	787-796
	797-806
	807-816
	817-826
	827-836
	837-846
	847-856
	857-866
	867-876
	877-886
	887-896
	897-906
	907-916
	917-926
	927-936
	937-946
	947-956
	957-966
	967-976
	977-986
	987-996
	997-1006
	1007-1016
	1017-1026
	1027-1036
	1037-1046
	1047-1056
	1057-1066
	1067-1076
	1077-1086
	1087-1096
	1097-1106
	1107-1116
	1117-1126
	1127-1136
	1137-1146
	1147-1156
	1157-1166
	1167-1176
	1177-1186
	1187-1196
	1197-1206
	1207-1216
	1217-1226
	1227-1236
	1237-1246
	1247-1256
	1257-1266
	1267-1276
	1277-1286
	1287-1296
	1297-1306
	1307-1316
	1317-1326
	1327-1336
	1337-1346
	1347-1356
	1357-1366
	1367-1376
	1377-1386
	1387-1396
	1397-1406
	1407-1416
	1417-1426
	1427-1436
	1437-1446
	1447-1456
	1457-1466
	1467-1476
	1477-1486
	1487-1496
	1497-1506
	1507-1516
	1517-1526
	1527-1536
	1537-1546
	1547-1556
	1557-1566
	1567-1576
	1577-1586
	1587-1596
	1597-1606
	1607-1616
	1617-1626
	1627-1636
	1637-1646
	1647-1656
	1657-1666

AC012230
AC012230.3 GI:7637254
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens. clone Rp11-1M10

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

* * be preserved.	¹	1003:	contig of	1003 bp in length
*	1004	1103:	gap of	100 bp
*	1104	2634:	Contig of	1531 bp in length
*	2635	2734:	gap of	100 bp
*	2735	4415:	Contig of	1681 bp in length
*	4416	4515:	gap of	100 bp
*	4516	5785:	Contig of	1270 bp in length
*	5786	5885:	gap of	100 bp
*	5886	7879:	Contig of	1994 bp in length
*	7880	9799:	gap of	100 bp
*	7980	9686:	Contig of	1707 bp in length
*	9687	9786:	gap of	100 bp

Query Match	10.8%	Score 348;	DB 2;	Length 69570;
Best Local Similarity	99.7%;	Pred. No. 7.2e-130;		
Matches 398;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;
*	51505	52372:	contig of 868 bp in length	
*	52373	52472:	gap of 100 bp	
*	52473	53328:	contig of 856 bp in length	
*	53329	53428:	gap of 100 bp	
*	53429	54268:	contig of 840 bp in length	
*	54269	54368:	gap of 100 bp	
*	54369	55229:	contig of 861 bp in length	
*	55230	55329:	gap of 100 bp	
*	55330	56197:	contig of 868 bp in length	
*	56198	56297:	gap of 100 bp	
*	56298	57163:	contig of 866 bp in length	
*	57164	57263:	gap of 100 bp	
*	57264	58130:	contig of 867 bp in length	
*	58131	58230:	gap of 100 bp	
*	58231	59082:	contig of 852 bp in length	
*	59083	59182:	gap of 100 bp	
*	59183	60020:	contig of 838 bp in length	
*	60021	60120:	gap of 100 bp	
*	60121	60983:	contig of 863 bp in length	
*	60984	61083:	gap of 100 bp	
*	61084	61935:	contig of 852 bp in length	
*	61936	62035:	gap of 100 bp	
*	62036	62866:	contig of 831 bp in length	
*	62867	62966:	gap of 100 bp	
*	62967	63827:	contig of 861 bp in length	
*	63828	63927:	gap of 100 bp	
*	63928	64783:	contig of 856 bp in length	
*	64784	64883:	gap of 100 bp	
*	64884	65740:	contig of 857 bp in length	
*	65741	65840:	gap of 100 bp	
*	65841	66684:	contig of 844 bp in length	
*	66685	66784:	gap of 100 bp	
*	66785	67653:	contig of 867 bp in length	
*	67652	67751:	gap of 100 bp	

RESULT	13				
AC012230/c					
LOCUS	AC012230	175064 bp	DNA	linear	HTG 22-APR-2000
DEFINITION	Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered pieces.				

* * * * *	9787 12254 12354 15329 15329 17201 17301 20232 22587 22588 25708 25808 28185 28285 31339 31439 31439 34300 34300 38319 38418 38419 42836 42935 42935 45449 45449 45449 48117 48216 52619 52718 52719 56593 56693 59635 59736 63661 63662 63762 68438 68538 71459 71559 76889 76989 82114 82214 88321 88321 93500 93500 93600 97901 98002 98002 103017 103117 103117 109179 109279 117308 117408 124080 124179 124180 131282 131282 131382 138060 138159 145491 145592 157391 157392 157491 157492	12253: contig of 2467 bp in length 12353: gap of 100 bp 15288: contig of 2875 bp in length 15328: gap of 100 bp 17200: contig of 1872 bp in length 17300: gap of 100 bp 20131: contig of 2831 bp in length 20231: gap of 100 bp 22587: contig of 2356 bp in length 22687: gap of 100 bp 25707: contig of 3020 bp in length 25807: gap of 100 bp 28184: contig of 2377 bp in length 28284: gap of 100 bp 31338: contig of 3054 bp in length 31438: gap of 100 bp 34299: contig of 2861 bp in length 34399: gap of 100 bp 38318: contig of 3919 bp in length 38418: gap of 100 bp 42835: contig of 4417 bp in length 42935: gap of 100 bp 45448: contig of 2513 bp in length 45548: gap of 100 bp 48116: contig of 2568 bp in length 48216: gap of 100 bp 52618: contig of 4402 bp in length 52718: gap of 100 bp 55592: contig of 3874 bp in length 56692: gap of 100 bp 59635: contig of 2943 bp in length 59735: gap of 100 bp 63661: contig of 3926 bp in length 63761: gap of 100 bp 68437: contig of 4676 bp in length 68537: gap of 100 bp 71458: contig of 2921 bp in length 71558: gap of 100 bp 76888: contig of 5330 bp in length 76988: gap of 100 bp 82213: contig of 5125 bp in length 82214: gap of 100 bp 88220: contig of 6007 bp in length 88320: gap of 100 bp 93499: contig of 5179 bp in length 93500: gap of 100 bp 97901: contig of 4302 bp in length 98001: gap of 100 bp 103016: contig of 5015 bp in length 103117: gap of 100 bp 103178: contig of 6062 bp in length 109278: gap of 100 bp 117307: contig of 8029 bp in length 117407: gap of 100 bp 124079: contig of 6672 bp in length 124179: gap of 100 bp 131281: contig of 7102 bp in length 131381: gap of 100 bp 138059: contig of 6678 bp in length 138159: gap of 100 bp 145491: contig of 7332 bp in length 145591: gap of 100 bp 157391: contig of 11800 bp in length 157491: gap of 100 bp 175064: contig of 17573 bp in length.
FEATURES		
source	Location/Qualifiers	
	1. .175064	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="Rep1-1M10"	
misc_feature	/clone_lib="RPI-11 Human Maie BAC"	
	1. .1003	
	/note="assembly_fragment"	

```

|||||
Db 1405 TCCCAAGAGCTCAAGAGTTCAACAAGTTCACTTTTCAGAAACCCCTTTGAGGAAGAC 1346
|||||
Qy 335 agaataacatctctctccattttaagatgaagaacagcgccggcgacacattggctaat 394
|||||
Db 1345 AGAATATACATCTCTCTCCATTTTAAAGATGAAGAAACAGCGCGGCGCACAAATGGCTAAT 1286
|||||
Qy 395 gctgttaatccacgaactttggagggtgaggcca 429
|||||
Db 1285 GCCTGTATCCAGCACTTTGGGAGGCTGAGGCCA 1251
|||||

RESULT 14
AK022254 1750 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens cDNA FLJ12192 fis, clone MAMMAL000851.
DEFINITION AK022254
ACCESSION AK022254.1 GI:10433612
VERSION Homo sapiens Mammary gland cDNA to mRNA, clone_lib:MAMMAL
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE clone:MAMMAL000851.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,K., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahara,K., Masubo,Y. and Oshima,A.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 1750)
AUTHORS Isogai,T. and Otsuki,T.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
source
1..1750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMAL000851"
/clone_lib="MAMMAL"
/tissue_type="Mammary gland"
/note="cloning vector: pME18SFL3"
BASE COUNT 291 a 489 c 586 g 384 t
ORIGIN

Query Match 10.4%; Score 336; DB 9; Length 1750;
Best Local Similarity 100.0%; Pred. No. 4.6e-183;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2896 aattcgagcagagtgagtgagggcgagccgagcagagcagagccctctctccc 2955
|||||
Db 1 AATTGCGAGCAGAGTGAGTGGGCGGGACCCGCGAGAGCGCGACCTTCTCTCCC 60
|||||
Qy 2956 gggctgcgcagggcagggcgaggagctccgcgcaccacacagagccggttctcaggggc 3015
|||||
Db 61 GGGCTGCGCAGGGCAGGGCGGGAGCTCCGGCACCACAGAGCGGTTCTCAGGGCGC 120
|||||
Qy 3016 ttgtctctgtttttcccggttctgttttctccctctccgaaggttgcagg 3075
|||||
Db 121 TTTGTCTCTGTTGTTTTCCTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGTTGTCAAGG 180
|||||

```

```

Qy 3076 ggtaggagaaagagagcgcaaacacaaagtgaagagctctccagtact 3135
|||||
Db 181 GGTAGGAGAAAGAGAGCGCAACACAAAAGTGGAAACAGGTAGAGGCTCTCCAGTGACT 240
|||||
Qy 3136 tacttggcggtattattttgttttcgagggcgaagggcttcgggaagtgcctcggttcg 3195
|||||
Db 241 TACTTGGCGGTATTGTTTGTTCGAGGCCAAGGAGGCTTCGGGAAGTGCTCGGTTTCG 300
|||||
Qy 3196 gggactttgatccggagccccacatccccaccactt 3231
|||||
Db 301 GGGACTTTGATCGGAGCCGCCACATCCCACTT 336
|||||

RESULT 15
AF258627 697 bp mRNA linear PRI 11-MAY-2000
LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
DEFINITION partial cds.
ACCESSION AF258627
VERSION AF258627.1 GI:7769707
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Analysis of hABCI gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
FEATURES
Location/Qualifiers
source
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/note="ABC1"
396..>697
/gene="ABCA1"
/note="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AAF69513.1"
/db_xref="GI:7769708"
/translation="MACWFLRLLLKMLNLTERRRQTCQLLELVAVPLFLILISVRL
SYPPVEQHECHPNKAPSGATLPWVGIIICNANNPCFRYPPTGPEAGVVGFNKS"
BASE COUNT 152 a 198 c 190 g 156 t 1 others
ORIGIN

Query Match 9.4%; Score 303; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 6.1e-164;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2812 agagcacaggtttgaccgatagtaacctctgcctcggtgcagccgaatataaagg 2871
|||||
Db 1 AGACACAGAGGCTTTGACCGATAGTAACCTCTCGCTCGTGCAGCCGAATATATAAAGG 60
|||||
Qy 2872 aactagtcgccgcaaaaaccccgtaattgcgagcgagagtgagtgggcgggaccgcga 2931
|||||
Db 61 AACTAGTCCCGCAAAAACCCCGTAATTGCGAGCGAGTGAGTGGGCGCGGACCCGCA 120
|||||

```

QY 2932 gagccgagccagccttctctccgggctcggcaggcagggcagggagctccgcgcac 2991
 |||
 Db 121 GAGCCGAGCGGACCTTCTCTCCGGGCTCGGCGAGGCGGGAGCTCCGGGCAC 180
 |||
 QY 2992 caacagagccggtctcagggcgcttgcctctgttttcccccggtctgtttctcc 3051
 |||
 Db 181 CAACAGAGCGGTTCTCAGGCGCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTCTCC 240
 |||
 QY 3052 cctctccgaagccttctcaaggggttagaagaagacgcaaacacaaaagtgaataa 3111
 |||
 Db 241 CTTTCTCCGAAGCCTTGTCAAGGGGTGAGGAAGAGACGCAACACAAAAGTGAATAA 300
 |||
 QY 3112 cag 3114
 |||
 Db 301 CAG 303

RESULT 16

AC021345/c

LOCUS

DEFINITION

AC021345

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC021345
 Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
 AC021345
 AC021345.2 GI:9130845
 HTG: HTGS_PHASE0.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 90698)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Anderson, S., Balgwin, J., Barna, N., Beckerly, R., Beda, F.,
 Boguslavsky, L., Bouckgater, B., Brown, A., Burkett, G., Castle, A.,
 Choepli, J., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
 DeArrellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
 Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
 Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lehotzky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheters, R., Meldrim, J., Menes, L., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6705761.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4483
 Center clone name: 24_J9

* NOTE: This record contains 92 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 910: contig of 910 bp in length
 911 1010: gap of 100 bp
 1011 1873: contig of 863 bp in length
 1874 1973: gap of 100 bp
 1974 2824: contig of 851 bp in length
 2825 2924: gap of 100 bp
 2925 3802: contig of 878 bp in length
 3803 3902: gap of 100 bp
 3903 4816: contig of 914 bp in length
 4817 4916: gap of 100 bp
 4917 5759: contig of 843 bp in length
 5760 5859: gap of 100 bp
 5860 6764: contig of 905 bp in length
 6765 6864: gap of 100 bp
 6865 7747: contig of 883 bp in length
 7748 7847: gap of 100 bp
 7848 8755: contig of 908 bp in length
 8756 8855: gap of 100 bp
 8856 9753: contig of 898 bp in length
 9754 9853: gap of 100 bp
 9854 10757: contig of 904 bp in length
 10758 10857: gap of 100 bp
 10858 11732: contig of 875 bp in length
 11733 11832: gap of 100 bp
 11833 12739: contig of 907 bp in length
 12740 12839: gap of 100 bp
 12840 13710: contig of 871 bp in length
 13711 13810: gap of 100 bp
 13811 14684: contig of 874 bp in length
 14685 14784: gap of 100 bp
 14785 15662: contig of 878 bp in length
 15663 15762: gap of 100 bp
 15763 16677: contig of 915 bp in length
 16678 16777: gap of 100 bp
 16779 17678: contig of 901 bp in length
 17679 17778: gap of 100 bp
 17779 18678: contig of 901 bp in length
 18680 18778: gap of 100 bp
 18780 19632: contig of 853 bp in length
 19633 19732: gap of 100 bp
 19733 20634: contig of 902 bp in length
 20635 20734: gap of 100 bp
 20735 21620: contig of 886 bp in length
 21621 21720: gap of 100 bp
 21721 22579: contig of 859 bp in length
 22580 22679: gap of 100 bp
 22680 23568: contig of 889 bp in length
 23569 23668: gap of 100 bp
 23669 24554: contig of 886 bp in length
 24555 24654: gap of 100 bp
 24655 25521: contig of 867 bp in length
 25522 25621: gap of 100 bp
 25622 26487: contig of 866 bp in length
 26488 26587: gap of 100 bp
 26588 27464: contig of 877 bp in length
 27465 27564: gap of 100 bp
 27565 28466: contig of 902 bp in length
 28467 28566: gap of 100 bp
 28567 29464: contig of 898 bp in length
 29465 29564: gap of 100 bp
 29565 30447: contig of 883 bp in length
 30448 30547: gap of 100 bp
 30548 31453: contig of 906 bp in length
 31454 31553: gap of 100 bp
 31554 32452: contig of 899 bp in length
 32453 32552: gap of 100 bp
 32553 33447: contig of 895 bp in length
 33448 33547: gap of 100 bp

JOURNAL diagnostic application
Patent: WO 0130846-A 70 03-MAY-2001;
Aventis Pharma S.A. (FR)

FEATURES
source
1..9854
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2665 a 2219 c 2334 g 2635 t 1 others
ORIGIN

Query Match 6.3%; Score 205; DB 6; Length 9854;
Best Local Similarity 100.0%; Pred. No. 6.7e-107;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtgggcccggaccgagcgagccgagccgagcccttctccgggctcgccgaggg 2969
Db 1 GTGAGTGGGGCCCGGACCGGAGCGGAGCGGACCCCTTCTCCGGGCTCGCGCAGGG 60
QY 2970 caggcgaggagctccgacacacagacgagccggttctcaggcgcttctgctctgttt 3029
Db 61 CAGGGCGGGGAGCTCCGGCCGACCAAGAGCCGGTCTCAGGGCGCTTTCCTCTGTTT 120
QY 3030 ttccccgggtctgttttctcccttctccggaagcgttgcaggggtagagaaagag 3089
Db 121 TTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAAGGGGTAGGAGAAAGAG 180
QY 3090 acgcaaacacaaagtggaaacag 3114
Db 181 ACGCAAAACAAAAGTGGAAACAG 205

RESULT 21
LOCUS AX139818 9854 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 70 from Patent EP1096012.
ACCESSION AX139818
VERSION AX139818.1 GI:14275400
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9854)
AUTHORS Denefle P., Rosier-Montus M.F., Arnould-Reguigne, I., Prades, C., Naudin, L., Lemoine, C., Duverger, N., Jave, M., searfof Iii, G.H., Remaley, A., Brewer, H.B. and Dean, M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and diagnostic application
JOURNAL Patent: EP 1096012-A 70 02-MAY-2001;
Aventis Pharma S.A. (FR)

FEATURES
source
1..9854
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2665 a 2219 c 2334 g 2635 t 1 others
ORIGIN

Query Match 6.3%; Score 205; DB 6; Length 9854;
Best Local Similarity 100.0%; Pred. No. 6.7e-107;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtgggcccggaccgagcgagccgagccgagcccttctccgggctcgccgaggg 2969
Db 1 GTGAGTGGGGCCCGGACCGGAGCGGAGCGGACCCCTTCTCCGGGCTCGCGCAGGG 60
QY 2970 caggcgaggagctccgacacacagacgagccggttctcaggcgcttctgctctgttt 3029
Db 61 CAGGGCGGGGAGCTCCGGCCGACCAAGAGCCGGTCTCAGGGCGCTTTCCTCTGTTT 120
QY 3030 ttccccgggtctgttttctcccttctccggaagcgttgcaggggtagagaaagag 3089
Db 181 ACGCAAAACAAAAGTGGAAACAG 205

Db 121 TTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAAGGGGTAGGAGAAAGAG 180
QY 3090 acgcaaacacaaagtggaaacag 3114
Db 181 ACGCAAAACAAAAGTGGAAACAG 205

RESULT 22
LOCUS AX060713 10442 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 1 from Patent WO0078972.
ACCESSION AX060713
VERSION AX060713.1 GI:12406103
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn, R.M., Wade, D. and Garvin, M.
TITLE Regulation with binding cassette transporter protein abcl
JOURNAL Patent: WO 0078972-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)

FEATURES
source
1..10442
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
ORIGIN

Query Match 6.1%; Score 197; DB 6; Length 10442;
Best Local Similarity 100.0%; Pred. No. 3e-102;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2918 ggcggggaccgagcgagccgagcccttctccgggctcgccgagggcggcgg 2977
Db 1 GGCGGGGACCGCGCAGAGCGCGCCGCCCTTCTCCGGGCTCGCGCAGGGCGGG 60
QY 2978 ggagctccgacacacagagccggttctcaggcgcttctgctctgtttttccccc 3037
Db 61 GGAGCTCCCGCGCCACCAACAGACCGGTTCTCAGGGCGCTTCTCTCTTTTCCCG 120
QY 3038 gttctgttttctcccttctccggaagcgttgcaggggttaggagaagagcgaac 3097
Db 121 GTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAAGGGGTAGGAGAAAGAGACGCAAC 180
QY 3098 acaaaagtggaaacag 3114
Db 181 ACAAAGTGGAAACAG 197

RESULT 23
LOCUS AX060892 10442 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 1 from Patent WO0078971.
ACCESSION AX060892
VERSION AX060892.1 GI:12406270
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE ATP binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)

FEATURES
source
1..10442
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others


```
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others
ORIGIN

Query Match 5.8%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 5e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccacccttctcccggtcgccgagcagcggcgagctccg 2986
Db 42 CCGCAGAGCCGAGCCGACCTTCTCTCCCGGCTGCGCAGGCGGCGGAGCTCCG 101
QY 2987 cgcaccaacagagccggttctcagggcgcttgcctctgtttttcccggtctgtt 3046
Db 102 CGCACCAACAGAGCCGCTTCTCAGGGCGCTTCTCTCTGTTTTCCTCGTTCGTTT 161
QY 3047 tctcccttctccgagggcttgcaggggttaggagagacgacacacaaagt 3106
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCAACAAAGTG 221
QY 3107 gaaaacag 3114
Db 222 GAAACAG 229

RESULT 26
AX060721
LOCUS AX060721 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078972.
ACCESSION AX060721
VERSION AX060721.1 GI:12406109
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE Regulation with binding cassette transporter protein abcl
JOURNAL Patent: WO 0078972-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
Location/Qualifiers
1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN

Query Match 5.8%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 5e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccacccttctcccggtcgccgagcagcggcgagctccg 2986
Db 42 CCGCAGAGCCGAGCCGACCTTCTCTCCCGGCTGCGCAGGCGGCGGAGCTCCG 101
QY 2987 cgcaccaacagagccggttctcagggcgcttgcctctgtttttcccggtctgtt 3046
Db 102 CGCACCAACAGAGCCGCTTCTCAGGGCGCTTCTCTCTGTTTTCCTCGTTCGTTT 161
QY 3047 tctcccttctccgagggcttgcaggggttaggagagacgacacacaaagt 3106
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCAACAAAGTG 221
QY 3107 gaaaacag 3114
Db 222 GAAACAG 229

RESULT 27
AX060900
LOCUS AX060900 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078971.
ACCESSION AX060900
VERSION AX060900.1 GI:12406276
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE Atp binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
Location/Qualifiers
1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN

Query Match 5.8%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 5e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccacccttctcccggtcgccgagcagcggcgagctccg 2986
Db 42 CCGCAGAGCCGAGCCGACCTTCTCTCCCGGCTGCGCAGGCGGCGGAGCTCCG 101
QY 2987 cgcaccaacagagccggttctcagggcgcttgcctctgtttttcccggtctgtt 3046
Db 102 CGCACCAACAGAGCCGCTTCTCAGGGCGCTTCTCTCTGTTTTCCTCGTTCGTTT 161
QY 3047 tctcccttctccgagggcttgcaggggttaggagagacgacacacaaagt 3106
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCAACAAAGTG 221
QY 3107 gaaaacag 3114
Db 222 GAAACAG 229
```

```

|||||
Db 42 CGCGAGAGCGAGCGACCTTCTCCGGGCTCGCGAGCGAGCGGAGCTCCG 101
QY 2987 cgcacacacagcgggttcagggcgttgccttgcctttttttcccggttctgtt 3046
|||||
Db 102 CGACCAACAGAGCGGTTCTCAGGGCGTTTCCTGTCCTGTTTTCCTCCGGTCTGTTT 161
QY 3047 tctcccttctccgaagcgttgcacgggtaggagaaagagacgcaaacacacaaagt 3106
|||||
Db 162 TCTCCCTCTCTCCGAGGCTGTCTCAGGGGTAGAGAAAGAGAGCGCAACACAAAGTG 221
QY 3107 gaaaacag 3114
|||||
Db 222 GAAACAG 229

RESULT 29
AC021345
LOCUS Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC021345 linear HTG 13-JUL-2000
ACCESSION AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 90698)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-24J9
Unpublished
REFERENCE 2 (bases 1 to 90698)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguski,K., Bouknight,B., Brown,A., Burkett,G., Castle,A.,
Choepell,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArrellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lander,T., Lechoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705761.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4483
Center clone name: 24_J9
-----
* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone

```

```

* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1873: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length
* 2825 2924: gap of 100 bp
* 2925 3802: contig of 878 bp in length
* 3803 3902: gap of 100 bp
* 3903 4816: contig of 914 bp in length
* 4817 4916: gap of 100 bp
* 4917 5759: contig of 843 bp in length
* 5760 5859: gap of 100 bp
* 5860 6764: contig of 905 bp in length
* 6765 6864: gap of 100 bp
* 6865 7747: contig of 883 bp in length
* 7748 7847: gap of 100 bp
* 7848 8755: contig of 908 bp in length
* 8756 8855: gap of 100 bp
* 8856 9753: contig of 898 bp in length
* 9754 9853: gap of 100 bp
* 9854 10757: contig of 904 bp in length
* 10758 10857: gap of 100 bp
* 10858 11732: contig of 875 bp in length
* 11733 11832: gap of 100 bp
* 11833 12739: contig of 907 bp in length
* 12740 12839: gap of 100 bp
* 12840 13710: contig of 871 bp in length
* 13711 13810: gap of 100 bp
* 13811 14684: contig of 874 bp in length
* 14685 14784: gap of 100 bp
* 14785 15662: contig of 878 bp in length
* 15663 15762: gap of 100 bp
* 15678 16677: contig of 915 bp in length
* 16678 16777: gap of 100 bp
* 16779 17778: contig of 901 bp in length
* 17779 18679: contig of 901 bp in length
* 18680 18779: gap of 100 bp
* 18780 19632: contig of 853 bp in length
* 19633 19732: gap of 100 bp
* 19733 20634: contig of 902 bp in length
* 20635 20734: gap of 100 bp
* 20735 21620: contig of 886 bp in length
* 21621 21720: gap of 100 bp
* 21721 22579: contig of 859 bp in length
* 22580 22679: gap of 100 bp
* 22680 23568: contig of 889 bp in length
* 23569 23668: gap of 100 bp
* 23669 24554: contig of 886 bp in length
* 24555 24654: gap of 100 bp
* 24655 25521: contig of 867 bp in length
* 25522 25621: gap of 100 bp
* 25622 26487: contig of 866 bp in length
* 26488 26587: gap of 100 bp
* 26588 27464: contig of 877 bp in length
* 27465 27564: gap of 100 bp
* 27565 28466: contig of 902 bp in length
* 28467 28566: gap of 100 bp
* 28567 29464: contig of 898 bp in length
* 29465 29564: gap of 100 bp
* 29565 30447: contig of 883 bp in length
* 30448 30547: gap of 100 bp
* 30548 31453: contig of 906 bp in length
* 31454 31553: gap of 100 bp
* 31554 32452: contig of 899 bp in length
* 32453 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp

```

* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length
* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
* 54226 54325: gap of 100 bp
* 54326 55206: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length
* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match 4.7%; Score 152; DB 2; Length 90698;
Best Local Similarity 99.2%; Pred. No. 5,1e-76;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 35 tactcggagggtggaggttgcaatgagccagatgcacattgcactccagcctgggca 94
|||||
Db 10177 TACTCGGAGGTGGAGGTGCAATGAGCCAGATGCACATTGCACCTCAGCCTGGCA 10236
|||||
QY 95 acaaaagtgaaactccatctcaattataaaaaaagaaagatttggtggtcgacttca 154
|||||
Db 10237 ACAAAGGTGAACTCCATCTCAATTAATAAAAAATAATGATTTGGTGGTCGACTTCA 10296
|||||
QY 155 aatagtaggagaagaagagagagatgaggtgcaggagatcttaattactctcta 214
|||||
Db 10297 AATAGTAGGAGAAGAAGGAGAGAGATGGAGGTACAGGAGATCTAATTACTCTCTA 10356
|||||
QY 215 aatcatgctaggaagataaacaccttttaataaactctctgcttttataacatactc 274
|||||
Db 10357 AATCATGCTAGGAAGATAACACCTTTTAATAACACTCTCTGCTTTTATAACATCATTC 10416
|||||
QY 275 tgccaagagactca 288
|||||
Db 10417 TGCCAAGGAGCTCA 10430
|||||

RESULT 30
AX351034 AX351034 117 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 6 from Patent WO0183746.
DEFINITION AX351034
ACCESSION AX351034
VERSION AX351034.1 GI:18616390
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 6 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1. .117
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 20 a 30 c 34 g 33 t
ORIGIN

Query Match 3.6%; Score 117; DB 6; Length 117;
Best Local Similarity 100.0%; Pred. No. 5,4e-56;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3115 gtaagaggctctccagtgacttacttggtggcggtattgtttgttcgagcccaaggagc 3174
|||||
Db 1 GTAAGAGGCTCTCCAGTGACTTACTTGGGCGTTATTGTTTGTTCGAGCCCAAGGAGGC 60
|||||
QY 3175 ttcgggaagtgcctcggttttcggggactttgatccggagcccccaatccccaccatt 3231
|||||
Db 61 TTCGGGAAGTGCCTGGGTTTCGGGACTTTTGATCCGGAGCCCCACATCCCCACCATT 117
|||||

RESULT 31
AX127764 AX127764 446 bp DNA linear PAT 15-MAY-2001
LOCUS Sequence 3 from Patent WO0130848.
DEFINITION AX127764
ACCESSION AX127764
VERSION AX127764.1 GI:14134411
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (bases 1 to 446)
AUTHORS Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 3 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN
Query Match 2.8%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 2.1e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTCGAAGGGGTAGGA 60
QY 3083 gaaagagacgcaaacacaaagtggaaacag 3114
Db 61 GAAAGAGACGCAACACAAAGTGGAAAAACAG 92
RESULT 32
AXI39751
LOCUS AXI39751 446 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1096012.
ACCESSION AXI39751
VERSION AXI39751.1 GI:14275333
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 446)
AUTHORS Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss Iii,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN
Query Match 2.8%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 2.1e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTCGAAGGGGTAGGA 60
QY 3083 gaaagagacgcaaacacaaagtggaaacag 3114
Db 61 GAAAGAGACGCAACACAAAGTGGAAAAACAG 92
RESULT 33
AXI27830
LOCUS AXI27830 9741 bp DNA linear PAT 15-MAY-2001

DEFINITION Sequence 69 from Patent WO0130848.
ACCESSION AXI27830
VERSION AXI27830.1 GI:14134477
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9741)
AUTHORS Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 69 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 2.8%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 2.9e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTCGAAGGGGTAGGA 60
QY 3083 gaaagagacgcaaacacaaagtggaaacag 3114
Db 61 GAAAGAGACGCAACACAAAGTGGAAAAACAG 92
RESULT 34
AXI39817
LOCUS AXI39817 9741 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 69 from Patent EP1096012.
ACCESSION AXI39817
VERSION AXI39817.1 GI:14275399
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9741)
AUTHORS Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss Iii,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 69 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 2.8%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 2.9e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTCGAAGGGGTAGGA 60
QY 3083 gaaagagacgcaaacacaaagtggaaacag 3114

```

|||||
Db 61 GAAAGAGCGCAACACAAAAGTGGAACAG 92
|||||
RESULT 35
AX351038
LOCUS AX351038
DEFINITION Sequence 10 from Patent WO0183746.
ACCESSION AX351038
VERSION AX351038.1 GI:18616393
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Densfle, P.,
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl1 gene
JOURNAL Patent: WO 0183746-A 10 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
Location/Qualifiers
1. .9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 2.8%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 2.9e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3023 cttgtttttcccggttctgttttcccttccttcggaaggtttcgaagggttagga 3082
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCCTCCCTTCCTCCGGAAGGCTTCTCAAGGGGTAGGA 60
|||||
Qy 3083 gaaagagcgaacacacaaaagtggaaaacag 3114
|||||
Db 61 GAAAGAGCGCAACACAAAAGTGGAACAG 92
|||||
RESULT 36
AC004139
LOCUS AC004139
DEFINITION Homo sapiens chromosome 17, clone HRP843B9, complete sequence.
ACCESSION AC004139
VERSION AC004139.1 GI:3513309
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 126295)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone HRP843B9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 126295)
AUTHORS Birren, B., Fasman, K., McKernan, K., Nusbaum, C., Richardson, P.,
Lander, E., Allen, N., Baker, J., Baldwin, J., Barna, N., Beckerly, R.,
Boutwell, C., Byrne, S., Cantu, C., Castle, A., Cerny, J., Cooke, P.,
Daly, M., Depayre, E., Devon, K., Dewar, K., Donelan, L., DuRette, B.,
Etamadi, S., Ferreira, P., Forrest, C., Funke, R., Gage, D., Gardyna, S.,
Gensheimer, S., Geraigery, K., Gilmartin, T., Gray, D., Hagos, B.,
Harris, K., Horton, L., Howland, J.C., Hui, L., Jacotot, L., Linton, L.,
MacKenzie, J., Marquis, N., McEwan, P., McGurk, A., Meldrim, J.,
Molla, M., Morris, W., Morrow, J., Nachman, A., Naylor, J., O'Connor, T.,
Paylin, B., Peterson, K., Ranganath, S., Riley, R., Roberts, D.,
Rollins, G., Rossello, R., Roy, A., Shyam, R., Soohoo, S.,
Stange-Thomann, N., Stilwell, J., Stone, C., Strickland, C., Sydney, K.,
Tang, L., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y.,
Ye, W.J., Zemtseva, I., Zhao, J. and Zody, M.
Direct Submission

```

```

JOURNAL Submitted (13-FEB-1998) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 126295)
Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J.,
Boutwell, C., Brown, A., Castle, A., Cerny, J., Collangelo, M.,
Collins, S., Collumore, A., Cooke, P., Corliss, D., Depayre, E.,
Devon, K., Dewar, K., Donelan, L., Ferreira, P., Fitzhugh, W.,
Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G.,
Hagos, B., Hearford, A., Herena, L., Horton, L., Howland, J.C.,
Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,
Paylin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,
Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,
Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,
Ye, W.J., Zhao, J. and Zody, M.
Direct Submission
JOURNAL Submitted (02-SEP-1998) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS On Sep 2, 1998 this sequence version replaced gi:3451370.
COMMENT All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES
source
Location/Qualifiers
1. .126295
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="HRPC843B9"
/repeat_region complement(2..193)
repeat_region /rpt_family="AluX"
repeat_region complement(235..410)
repeat_region /rpt_family="MER5A"
repeat_region complement(479..640)
repeat_region /rpt_family="MIR"
repeat_region complement(714..1012)
repeat_region /rpt_family="AluSc"
repeat_region 1140..1269
repeat_region /rpt_family="L2"
repeat_region complement(2503..2646)
repeat_region /rpt_family="MER5B"
repeat_region complement(2647..2954)
repeat_region /rpt_family="AluJb"
repeat_region complement(2955..2998)
repeat_region /rpt_family="MER5B"
repeat_region 3758..4544
repeat_region /rpt_family="L2"
repeat_region 4743..4873
repeat_region /rpt_family="FLAM_C"
repeat_region complement(5068..5364)
repeat_region /rpt_family="AluJb"
repeat_region complement(6403..6675)
repeat_region /rpt_family="AluSc"
repeat_region 6782..6950
repeat_region /rpt_family="MIR"
repeat_region complement(7379..7734)
repeat_region /rpt_family="Tigger1"
repeat_region complement(7735..7840)
repeat_region /rpt_family="(GGGA)n"
repeat_region complement(7841..7876)
repeat_region /rpt_family="purine-rich"
repeat_region complement(7883..7988)
repeat_region /rpt_family="Tigger1"
repeat_region complement(7989..8597)
repeat_region /rpt_family="Tigger1"
repeat_region complement(8604..8741)
repeat_region /rpt_family="FLAM_C"
repeat_region complement(8759..9456)

```

```

repeat_region      /rpt_family="Tigger1"
                    complement(9467..9759)
repeat_region      /rpt_family="AluSx"
                    complement(9820..10276)
repeat_region      /rpt_family="Tigger1"
                    10754..11095
repeat_region      /rpt_family="MER47A"
                    complement(11106..11949)
repeat_region      /rpt_family="L2"
                    12033..12147
repeat_region      /rpt_family="L2"
                    12148..12235
repeat_region      /rpt_family="L2"
                    complement(13066..13222)
repeat_region      /rpt_family="L2"
                    13224..13251
repeat_region      /rpt_family="L2"
                    13252..13544
repeat_region      /rpt_family="L2"
                    13545..13769
repeat_region      /rpt_family="L2"
                    14208..14364
repeat_region      /rpt_family="L2"
                    complement(14835..15195)
repeat_region      /rpt_family="L2"
                    complement(15353..15497)
repeat_region      /rpt_family="L2"
                    complement(15662..15828)
repeat_region      /rpt_family="L2"
                    complement(16301..16607)
repeat_region      /rpt_family="L2"
                    17500..17802
repeat_region      /rpt_family="L2"
                    complement(18474..18774)
repeat_region      /rpt_family="L2"
                    18825..18902
repeat_region      /rpt_family="L2"
                    complement(19656..19771)
repeat_region      /rpt_family="L2"
                    19834..20129
repeat_region      /rpt_family="L2"
                    complement(20145..20168)
repeat_region      /rpt_family="L2"
                    complement(20169..20451)
repeat_region      /rpt_family="L2"
                    complement(20456..20765)
repeat_region      /rpt_family="L2"
                    complement(21023..21343)
repeat_region      /rpt_family="L2"
                    21344..21630
repeat_region      /rpt_family="L2"
                    complement(21631..21700)
repeat_region      /rpt_family="L2"
                    22291..22526
repeat_region      /rpt_family="L2"
                    23606..23899
repeat_region      /rpt_family="L2"
                    complement(24202..25601)
repeat_region      /rpt_family="L2"
                    complement(25816..25978)
repeat_region      /rpt_family="L2"
                    26161..26465
repeat_region      /rpt_family="L2"
                    complement(26630..26725)
repeat_region      /rpt_family="L2"
                    complement(26902..26958)
repeat_region      /rpt_family="L2"
                    complement(26959..27179)
repeat_region      /rpt_family="L2"
                    complement(27189..27489)
repeat_region      /rpt_family="L2"
                    27712..28009
repeat_region      /rpt_family="L2"
                    complement(28773..29080)
repeat_region      /rpt_family="L2"
                    29132..29580
repeat_region      /rpt_family="L2"
                    29962..30097
repeat_region      /rpt_family="L2"
                    30275..30566
repeat_region      /rpt_family="L2"
                    complement(31041..31353)
repeat_region      /rpt_family="L2"
                    complement(32362..32571)
repeat_region      /rpt_family="L2"
                    complement(33470..33781)
repeat_region      /rpt_family="L2"
                    complement(35504..35774)
repeat_region      /rpt_family="L2"
                    35830..35906
repeat_region      /rpt_family="L2"
                    35907..36194
repeat_region      /rpt_family="L2"
                    36204..36516
repeat_region      /rpt_family="L2"
                    36517..36682
repeat_region      /rpt_family="L2"
                    complement(36689..37629)

Query Match      2.0% Score 65; DB 9; Length 126295;
Best Local Similarity 100.0%; Pred. No. 1.8e-25;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1017 gatctcgactcaacttaacctctgctcccggttcaagcgattctctgcctcgcctc 1076
Db 31115 GATCTCGACTCACTGTAACCTCTGCTCCCGGTTCAAGCGATTCTCTGCCTCAGCCTC 31174

QY 1077 ctgag 1081
Db 31175 CTGAG 31179

RESULT 37
AC021850/c
LOCUS AC021850 206966 bp DNA linear PRI 09-MAY-2001
DEFINITION Homo sapiens BAC clone RP11-389E17 from 4, complete sequence.
ACCESSION AC021850
VERSION AC021850.8 GI:12863232
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 206966)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2. (bases 1 to 206966)
AUTHORS Kang, K., Maupin, R. and Ureta, M.
TITLE The sequence of Homo sapiens BAC clone RP11-389E17
JOURNAL Unpublished
REFERENCE 3. (bases 1 to 206966)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4. (bases 1 to 206966)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5. (bases 1 to 206966)
AUTHORS Waterston, R.

```

TITLE Direct Submission
 JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Feb 16, 2001 this sequence version replaced gi:111761497.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: saplens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0389E17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-62B4. Actual start of this clone is at base position 1 of RP11-389E17; actual end is at base position 206966 of RP11-389E17.

FEATURES

source	Location/Qualifiers
	1. 206966
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-389E17"
	/db_lib="RP11-11"
	3. 256
repeat_region	/rpt_family="L1"
repeat_region	389..509
repeat_region	/rpt_family="L1"
repeat_region	523..1023
repeat_region	/rpt_family="L1"
repeat_region	1088..1319
repeat_region	/rpt_family="L1"
repeat_region	1779..1908
repeat_region	/rpt_family="MIR"
repeat_region	1930..1950
repeat_region	/rpt_family="AT_rich"
repeat_region	2531..2753
repeat_region	/rpt_family="Alu"
repeat_region	2772..2792
repeat_region	/rpt_family="AT_rich"
repeat_region	2987..3022
repeat_region	/rpt_family="AT_rich"
repeat_region	3223..3260
	/rpt_family="L1"
	/rpt_family="AT_rich"
	3386..3523
	/rpt_family="MIR"
	3775..4559
	/rpt_family="ERV1"
	4595..4763
	/rpt_family="ERV1"
	5163..5193
	/rpt_family="(CA)n"
	5276..5297
	/rpt_family="AT_rich"
	7864..8032
	/rpt_family="MIR"
	8397..8660
	/rpt_family="BRVL"
	8877..8991
	/rpt_family="GA-rich"
	9049..9235
	/rpt_family="L2"
	9284..9364
	/rpt_family="L1"
	9388..9734
	/rpt_family="L1"
	11490..11628
	/rpt_family="L2"
	12620..12715
	/rpt_family="MER2_type"
	12726..12923
	/rpt_family="MaLR"
	12950..13462
	/rpt_family="MER2_type"
	17273..17508
	/rpt_family="Alu"
	18160..18215
	/rpt_family="L2"
	18689..18810
	/rpt_family="MIR"
	19217..19324
	/rpt_family="L1"
	20012..20292
	/rpt_family="L1"
	20476..20830
	/rpt_family="MaLR"
	23028..23085
	/rpt_family="AT_rich"
	23483..23612
	/rpt_family="Achoho"
	23878..23976
	/rpt_family="AT_rich"
	24263..24404
	/rpt_family="MIR"
	24706..24903
	/rpt_family="MIR"
	26677..27207
	/rpt_family="MER1_type"
	27597..27619
	/rpt_family="AT_rich"
	28709..28781
	/rpt_family="AT_rich"
	29420..29635
	/rpt_family="Alu"
	29865..29898
	/rpt_family="tRNA-Cys-TGC"
	29979..30116
	/rpt_family="MIR"
	30773..31036
	/rpt_family="L2"
	31039..31722
	/rpt_family="L2"
	32160..32333
	/rpt_family="CRI"
	33594..34090
	/rpt_family="L1"

```

repeat_region 34248..34275
/rpt_family="AT_rich"
repeat_region 34635..34825
/rpt_family="L1"
repeat_region 35013..35078
/rpt_family="AT_rich"
repeat_region 35645..35691
/rpt_family="(CA)n"
repeat_region 36784..37028
/rpt_family="L1"
repeat_region 37100..37171
/rpt_family="L1"
repeat_region 37381..37431
/rpt_family="AT_rich"
repeat_region 37582..38445
/rpt_family="ERVL"
repeat_region 39128..39148
/rpt_family="AT_rich"
repeat_region 40735..40755
/rpt_family="AT_rich"
repeat_region 40826..41030
/rpt_family="L1"

Query Match
Best Local Similarity 100.0%; Pred. No. 4e-23; Length 206966;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1017 gatcgactcactgaacctgctccggttcaagcgattctcctgcctcagcctc 1076
|||||
Db 45729 GATCGACTCACTGTAACCTCTGCTCCGGGTTCAAGGATTCTCCGCTCAGCCTC 45670

QY 1077 c 1077
Db 45669 c 45669

RESULT 38
AF287263
LOCUS AF287263 278572 bp DNA linear ROD 23-APR-2001
DEFINITION Mus musculus ATP-binding cassette 1, sub-family A, member 1 (Abca1)
ACCESSION AF287263
VERSION AF287263.1 GI:11611824
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Qiu,Y., Cavelier,L., Chiu,S., Yang,X., Rubin,E., Rubin,E. and Cheng,J.F.
1 (bases 1 to 278572)
TITLE Human and mouse abca1 comparative sequencing and transgenesis
studies revealing novel regulatory sequences
JOURNAL Genomics 73 (1), 66-76 (2001)
MEDLINE 21251004
AUTHORS Qiu,Y., Cavelier,L., Chiu,S., Rubin,E. and Cheng,J.-F.
REFERENCE 2 (bases 1 to 278572)
SUBMIT Direct Submission
Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
CA 94720, USA
FEATURES
Source
1..278572
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="BAC 129K10; 142M4; 197F14"
join(87184..87402,103551..103699,115064..115187,
119217..119358,120364..120482,133479..133600,
135582..135758,147083..147175,149791..150031,
152367..152506,152845..152961,156570..156767,
157459..157664,158200..158376,158821..159043,
160617..160838,161790..161994,163076..163189,
164646..164817,165738..165869,167171..167313,
167492..167629,168320..168540,169790..169862,
170735..170937,172636..172684,172855..172968,
173925..174073,175605..175729,176761..176859,
180181..180370,181179..181273,182395..182427,
183398..183503,191761..191829,192306..192475,
193631..193808,196619..196734,196845..196989,
198526..198649,200922..201051,201281..201401,
202114..202176,202890..202996,204955..205096,
205427..205558,206739..206842,207361..207453,
208135..208378,209638..209778)
/gene="Abca1"
/product="ATP-binding cassette 1, sub-family A, member 1"
87184..210783
/gene="Abca1"
join(119237..119358,120364..120482,133479..133600,
135582..135758,147083..147175,149791..150031,
152367..152506,152845..152961,156570..156767,
157459..157664,158200..158376,158821..159043,
160617..160838,161790..161994,163076..163189,
164646..164817,165738..165869,167171..167313,
167492..167629,168320..168540,169790..169862,
170735..170937,172636..172684,172855..172968,
173925..174073,175605..175729,176761..176859,
180181..180370,181179..181273,182395..182427,
183398..183503,191761..191829,192306..192475,
193631..193808,196619..196734,196845..196989,
198526..198649,200922..201051,201281..201401,
202114..202176,202890..202996,204955..205096,
205427..205558,206739..206842,207361..207453,
208135..208378,209638..209778)
/gene="Abca1"
/codon_start=1
/product="ATP-binding cassette 1, sub-family A, member 1"
/protein_id="AAG39073.1"
/db_xref="GI:11611825"
/translation="MPSAGTLPWVQGIICNANNPCRYPTPGAPGVGVGNFKSIYSR
LFSDAQLLYSQDTSIKDMHVRMLRQIKHPNSNKLQDFLVNFTSGFLQHLN
KLDAEVLRYNMDILKPVVTKLNSHTLPOTHLAETVLLDSGLAQELPSTKSK
SDMOEYMTNVSNSSTQIYQAVSRIVCGHGGGLKIKSLNVEONNYKALFSG
NTEEDVTFYDNSTTPYCNLMKNLESSPLSRILINKALKPLLVGLKILYPTPATRQ
VMAEVNKTVELAVFHDLEGMWELSPQIWTFMENSOEMLVRLTLDNRGNDWFQK
LDGLDWTADIMAFKAKNEDQVSPNGSVYTWREAFNETNQAIQTSRMCEVNLNKL
EPITVEVRLINKSMELLDERKWAGIVFTGTPDSVELPHVYKIKRMDIVNERNK
IKDGYWDPGRADPFEDMYWGVGFALQDVVEQAIIRVLTGSEKTKGVYVQOMPVC
YVDIDFURVMSRNPMTLAWISYVAVI IKSIYVEKEARLKETMRLGDLNGLWFS
WFVSSLIPLLVAGLLVILKLNLLPYSDPSVVFVLSVFMVTLIQCLFLSTLFSR
ANLAAGCGIITYTLPLVYLCVAMQDYVGFSGIKFASLSVPAFGGCEYFALFEEQ
GIGVQMDNLFESPVEEDGNLTATVASMULFDTELYGVMTWYIEAVPGQYIGPRPYF
PCTGSYWFGEIDEKSHPGSQSGVSEICMEEPHTLRLGVSTQNLVYRDMGMVAV
DGLALNFEGQITSTFGHNGAGKTTMSITGLPPTSGTAYILGDKDIRSEMSIQN
LGVCPOHNLFDMLTVVEHIFWYARLKLSEKHVKAEMEQALDGLPPLSKLSTQ
LSGSMOKLSVALAFVGGSKVILDEPTAGVDPSYRGIWELLKYLROGRTIILSTHH
MDEADILGDRALIIHSGKCCVSSSLFNQLGTYLILVKKDVSSSSCKNSST
VSLCKEDSVSSQSSDAGSDHSDTLTIDVSAISNLIRKHVSEARLVEDIGHETY
VLPYEAKEGAFVELFHEIDRSLDLSIGSYSETTLEEFILKVAESGVDAETSDG
TLPARNRRAFGDKQCLPFTEDDADPNDSIDIPESRETLLSGMDGKSGYQLKGW
KLTOQOFPVALLKRLIARRSRKGFQFVILPAVFCIALVFLSVIPRGKVPSELQ
PMWNYQYTFVSDNAPEDMGTOELLNALTQDGFGRCMENPILPTCLAGEEDWTI
SPVQSIDVLQNGNTMKNPACQSSDKIKMLPVCPGPGAGGLPPQPKPQADAI
LQNLGRNIDSLVKTIVYQLIAKSLKNKIWNFEYRGYSGLVSNVLPSPSHEVNDI
IKOMKLLKLITKSADRLSLGRFMAGLDTKNNKVMFNKNKWHAISSFLNNINAIL
RANLQGENPSQYGTAFNHLNLTQLQSEVALMTSVYVVPATLVIIIFICQOKSYV
VFLPIBRYSKAKHLQFISGVFVIYVSNVMDMNYVVPATLVIIIFICQOKSYV
SSTNLVLAIIILLLXGWSITPLWYPAFSEFKIPSTAVVLTSVNLFINGSVATFVL
ELFTNNKLNIDILKSFLIIPHECLGRLIDMKVKNQAMADALERGNRRVSPLSW
DLVGRNLFAWEGVVFVLTITVLIQYRFRPVKAKLPPLNDEDEDVRRERQRLD
GGGNDILTEKLTXYRRKRPADVRCIGIPPGCEGLLGVNAGKSGTTFKMTGTD
TPVTRGDALFNKLSILNHEVQNNQYCPQFADITELLTGREHVEFFALLRGVPEKE
VGKGEWAIRKGLVYKGYKASNYSGNKRKLSTAMALIGGPVPLVDEPTTGMGPKA
RRLFWNCALISVKEGSSVLTSHSMECEALCTRAIMVNGFRCLGSVQHLKNGFD
GYTIVYRIAGSNPDLKPVOEFGFLAPPGSVLKEKHNMQLQYLPSSLSLARIFSLLS
QSKRLHIEDIYSVSTLQDQVFNVAQSDDDHLKDLSLHKNQTVVDVAVSLQD

```



```

BASE COUNT      76801 a 58762 c 61256 g 81498 t 255 others
ORIGIN
EKKVESYV"
Query Match      1.9%: Score 61; DB 10; Length 278572;
Best Local Similarity 100.08; Pred. No. 4.1e-23;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2744 ctccacgtcttctgtgactgaactacataaacagagccgggaagggcgccgg 2803
|||||
Db 87033 CTCACGCTCTTCTGCTGAGTGACTGAACATACATAACAGAGCCGGGAAGGGCGCGG 87092
QY 2804 q 2804
Db 87093 g 87093

RESULT 39
AC107969/c
LOCUS
DEFINITION      Homo sapiens chromosome 11 clone CTD-2028015 map 11, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION      AC107969
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 57662)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone CTD-2028015
Unpublished
2 (bases 1 to 57662)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cook,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferrel,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Minova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I24518
Center clone name: 2028_O_15
-----
* NOTE: This record contains 73 individual

```

```

* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 717: contig of 717 bp in length
* 718 817: gap of 100 bp
* 818 1525: contig of 708 bp in length
* 1526 1625: gap of 100 bp
* 1626 2265: contig of 640 bp in length
* 2266 2365: gap of 100 bp
* 2366 3050: contig of 685 bp in length
* 3051 3150: gap of 100 bp
* 3151 3845: contig of 695 bp in length
* 3846 3945: gap of 100 bp
* 3946 4577: contig of 632 bp in length
* 4578 4677: gap of 100 bp
* 4678 5369: contig of 692 bp in length
* 5370 5469: gap of 100 bp
* 5470 6151: contig of 682 bp in length
* 6152 6251: gap of 100 bp
* 6252 6955: contig of 704 bp in length
* 6956 7055: gap of 100 bp
* 7056 7804: contig of 749 bp in length
* 7805 7904: gap of 100 bp
* 7905 8603: contig of 699 bp in length
* 8604 8703: gap of 100 bp
* 8704 9392: contig of 689 bp in length
* 9393 9492: gap of 100 bp
* 9493 10155: contig of 663 bp in length
* 10156 10255: gap of 100 bp
* 10256 10948: contig of 693 bp in length
* 10949 11048: gap of 100 bp
* 11049 11603: contig of 555 bp in length
* 11604 11703: gap of 100 bp
* 11704 12386: contig of 683 bp in length
* 12387 12486: gap of 100 bp
* 12487 13187: contig of 701 bp in length
* 13188 13287: gap of 100 bp
* 13288 13988: contig of 701 bp in length
* 13989 14088: gap of 100 bp
* 14089 14795: contig of 707 bp in length
* 14796 14895: gap of 100 bp
* 14896 15532: contig of 637 bp in length
* 15533 15632: gap of 100 bp
* 15633 16347: contig of 715 bp in length
* 16348 16447: gap of 100 bp
* 16448 17155: contig of 708 bp in length
* 17156 17255: gap of 100 bp
* 17256 17961: contig of 706 bp in length
* 17962 18061: gap of 100 bp
* 18062 18723: contig of 662 bp in length
* 18724 18823: gap of 100 bp
* 18824 19476: contig of 653 bp in length
* 19477 19576: gap of 100 bp
* 19577 20274: contig of 698 bp in length
* 20275 20374: gap of 100 bp
* 20375 21074: contig of 700 bp in length
* 21075 21174: gap of 100 bp
* 21175 21867: contig of 693 bp in length
* 21868 21967: gap of 100 bp
* 21968 22680: contig of 713 bp in length
* 22681 22780: gap of 100 bp
* 22781 23495: contig of 715 bp in length
* 23496 23595: gap of 100 bp
* 23596 24227: contig of 632 bp in length
* 24228 24327: gap of 100 bp
* 24328 25032: contig of 705 bp in length

```

* 25033 25132: gap of 100 bp
* 25133 25838: contig of 706 bp in length
* 25839 25938: gap of 100 bp
* 25939 26651: contig of 713 bp in length
* 26652 26751: gap of 100 bp
* 26752 27431: contig of 680 bp in length
* 27432 27531: gap of 100 bp
* 27532 28232: contig of 701 bp in length
* 28233 28332: gap of 100 bp
* 28333 29032: contig of 700 bp in length
* 29033 29132: gap of 100 bp
* 29133 29838: contig of 706 bp in length
* 29839 29938: gap of 100 bp
* 29939 30653: contig of 715 bp in length
* 30654 30753: gap of 100 bp
* 30754 31466: contig of 713 bp in length
* 31467 31566: gap of 100 bp
* 31567 32272: contig of 706 bp in length
* 32273 32372: gap of 100 bp
* 32373 33009: contig of 637 bp in length
* 33010 33109: gap of 100 bp
* 33110 33827: contig of 718 bp in length
* 33828 33927: gap of 100 bp
* 33928 34646: contig of 719 bp in length
* 34647 34746: gap of 100 bp
* 34747 35451: contig of 705 bp in length
* 35452 35551: gap of 100 bp
* 35552 36244: contig of 693 bp in length
* 36245 36344: gap of 100 bp
* 36345 36993: contig of 649 bp in length
* 36994 37093: gap of 100 bp
* 37094 37777: contig of 684 bp in length
* 37778 37877: gap of 100 bp
* 37878 38587: contig of 710 bp in length
* 38588 38687: gap of 100 bp
* 38688 39375: contig of 688 bp in length
* 39376 39475: gap of 100 bp
* 39476 40179: contig of 704 bp in length
* 40180 40279: gap of 100 bp
* 40280 40986: contig of 707 bp in length
* 40987 41086: gap of 100 bp
* 41087 41749: contig of 663 bp in length
* 41750 41849: gap of 100 bp
* 41850 42551: contig of 702 bp in length
* 42552 42651: gap of 100 bp
* 42652 43363: contig of 712 bp in length
* 43364 43463: gap of 100 bp
* 43464 44144: contig of 681 bp in length
* 44145 44244: gap of 100 bp
* 44245 44939: contig of 695 bp in length
* 44940 45039: gap of 100 bp
* 45040 45740: contig of 701 bp in length
* 45741 45840: gap of 100 bp
* 45841 46513: contig of 673 bp in length
* 46514 46613: gap of 100 bp
* 46614 47307: contig of 694 bp in length
* 47308 47407: gap of 100 bp
* 47408 48097: contig of 690 bp in length
* 48098 48197: gap of 100 bp
* 48198 48884: contig of 687 bp in length
* 48885 48984: gap of 100 bp
* 48985 49700: contig of 716 bp in length
* 49701 49800: gap of 100 bp
* 49801 50484: contig of 684 bp in length
* 50485 50584: gap of 100 bp
* 50585 51288: contig of 704 bp in length
* 51289 51388: gap of 100 bp
* 51389 52107: contig of 719 bp in length
* 52108 52207: gap of 100 bp
* 52208 52909: contig of 702 bp in length
* 52910 53009: gap of 100 bp
* 53010 53665: contig of 656 bp in length
* 53666 53765: gap of 100 bp

* 53766 54444: contig of 679 bp in length

Query Match 1.8%; Score 58; DB 2; Length 57662;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattgctcccaagtctggtattacagcatgagccactgcgccagc 1244
|||||
DB 39017 CTGCCACCTTGGCTCCCAAAGTCTGGATTACAGGATGAGCCACTGCCCCAGC 38960
|||||

RESULT 40
HS931E15/5

LOCUS
DEFINITION Human DNA sequence from clone 931E15 on chromosome Xq25. Contains STSs, GSSs and genomic marker DXS8098, complete sequence.
ACCESSION AL023575
VERSION AL023575.1 GI:3618163
KEYWORDS HTG; DXS8098.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 81874)
Pavitt,R.
Direct Submission
Submitted (24-SEP-1998) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 18, 1998 this sequence version replaced gi:3550203.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known, annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>
931E15 is from the library RPC15 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>
IMPORTANT: This sequence is not the entire insert of clone 931E15. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
The true left end of 424J12 (282207) is at 36643 in this sequence.
The true right end of 506G2 (282213) is at 37431.
Location/Qualifiers
1. 81874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q25"
/clone="RP5-931E15"
/clone_lib="RPC1-5"
1. 201
/note="AluX repeat: matches 94. .294 of consensus"
complement(28. .531)
/note="AluX repeat: matches 292. .1 of consensus"
569. .687
/note="AluJo repeat: matches 11. .127 of consensus"
702. .1002
/note="AluSp repeat: matches 1. .302 of consensus"
1010. .1178
/note="FRAM repeat: matches 5. .166 of consensus"


```

repeat_region 63917..64021
/note="LMD repeat: matches 984..1089 of consensus"
misc_feature complement(64074..64543)
/note="match: GSS: Em:AQ885350"
misc_feature complement(66096..66577)
/note="match: GSS: Em:AQ611651"
misc_feature 69990..70342
/note="match: SPS: Em:G21603"
misc_feature complement(76094..76635)
/note="match: GSS: Em:AQ277578"
misc_feature complement(76184..76636)
/note="match: GSS: Em:AQ812553"
misc_feature complement(76460..76628)
/note="match: GSS: Em:AQ140871"
misc_feature complement(88581..89052)
/note="match: GSS: Em:AQ201907"
misc_feature 89057..89594
/note="match: GSS: Em:AQ543464"
91625..91990
/note="THE1C repeat: matches 1..371 of consensus"
92079..92237
/note="LIP47 repeat: matches 5949..6125 of consensus"
BASE COUNT 32991 a 18802 c 18896 g 27671 t
ORIGIN

Query Match 1.8%; Score 58; DB 9; Length 98360;
Best Local Similarity 100.0%; Pred. No. 2e-21; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

Qy 1187 ctgcccacttgctcccaagtgctgggtacagcagtcagccactgcgccagc 1244
|||||
Db 18635 CTGCCACCTTGGCTCCCAAGTGTGGGATACAGCATGACCCACTGCGCCAGC 18578

RESULT 42
AC107939/c
LOCUS AC107939 145178 bp DNA linear HTG 13-FEB-2002
DEFINITION Homo sapiens chromosome 11 clone RP11-4809 map 11, WORKING DRAFT
SEQUENCE, 3 ordered pieces.
ACCESSION AC107939
VERSION AC107939.2 GI:18653699
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145178)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-4809
Unpublished
2 (bases 1 to 145178)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Grand-Pierre,N.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenda,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

```

TITLE
JOURNAL

COMMENT

zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 13, 2002 this sequence version replaced gi:18308666.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24018
Center clone name: 48_O_9
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 144836 bases at least Q40
Consensus quality: 144904 bases at least Q30
Consensus quality: 144952 bases at least Q20
Insert size: 141000; agarose-fp
Insert size: 144978; sum-of-contents
Quality coverage: 18.7 in Q20 bases; agarose-fp
Quality coverage: 18.2 in Q20 bases; sum-of-contents

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 68019: contig of 68019 bp in length
* 68020 68119: gap of 100 bp
* 68120 140460: contig of 72341 bp in length
* 140461 140560: gap of 100 bp
* 140561 145178: contig of 4618 bp in length.
Location/Qualifiers
1. 145178
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-4809"
/clone_lib="RPC1-11 Human Male BAC"
1. 68019
/note="assembly_fragment"
clone_end:SP6
vector_side:left
68120..140460
/note="assembly_fragment"
140561..145178
/note="assembly_fragment"
clone_end:T7
vector_side:right
46626 a 28900 c 28355 g 41097 t 200 others
ORIGIN

FEATURES
Source

misc_feature

misc_feature

misc_feature

BASE COUNT
ORIGIN

Query Match 1.8%; Score 58; DB 2; Length 145178;
Best Local Similarity 100.0%; Pred. No. 2.1e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1187 ctgcccacattggctcccaagtgctgggtacagcagtcagccactgcgccagc 1244
|||||
Db 74827 CTGCCACCTTGGCTCCCAAGTGTGGGATACAGCATGACCCACTGCGCCAGC 74770

```

RESULT 43
AC015494
LOCUS
DEFINITION Homo sapiens clone RP11-21E12, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION AC015494.6 GI:14209771
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
1 (bases 1 to 146312)
TITLE Homo sapiens, clone RP11-21E12
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 146312)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgaiter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2001 this sequence version replaced gi:13443216.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3990
Center clone name: 21_E_12
----- Summary Statistics
Sequencing vector: M13: M7815; 3% of reads
Sequencing vector: Plasmid; n/a; 97% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 142529 bases at least Q40
Consensus quality: 144014 bases at least Q30
Consensus quality: 144580 bases at least Q20
Insert size: 128000; agarose-fp
Insert size: 145212; sum-of-contigs
Quality coverage: 9.4 in Q20 bases; agarose-fp
Quality coverage: 8.3 in Q20 bas.
NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*
* 1 1732: contig of 1732 bp in length
* 1733 1832: gap of 100 bp
* 1833 2945: contig of 1113 bp in length
* 2946 3045: gap of 100 bp
* 3046 4116: contig of 1071 bp in length
* 4117 4216: gap of 100 bp
* 4217 5757: contig of 1541 bp in length
*

```

```

* 5758 5857: gap of 100 bp
* 5858 7545: contig of 1688 bp in length
* 7546 7645: gap of 100 bp
* 7646 11527: contig of 3892 bp in length
* 11528 11627: gap of 100 bp
* 11628 55477: contig of 43850 bp in length
* 55478 55577: gap of 100 bp
* 55578 66767: contig of 11190 bp in length
* 66768 66867: gap of 100 bp
* 66868 83973: contig of 17106 bp in length
* 83974 84073: gap of 100 bp
* 84074 107174: contig of 23101 bp in length
* 107175 107274: gap of 100 bp
* 107275 137395: contig of 30121 bp in length
* 137396 137495: gap of 100 bp
* 137496 146312: contig of 8817 bp in length.
FEATURES
Location/Qualifiers
source
1. .146312
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RP11-21E12"
/clone_lib="RPCI-11 Human Male BAC"
misc_feature
1. .1732
/note="assembly_fragment"
clone_end:SP6
vector_side:left
1833. .2945
/note="assembly_fragment"
3046. .4116
/note="assembly_fragment"
4217. .5757
/note="assembly_fragment"
5858. .7545
/note="assembly_fragment"
7646. .11527
/note="assembly_fragment"
11628. .55477
/note="assembly_fragment"
55578. .66767
/note="assembly_fragment"
66868. .83973
/note="assembly_fragment"
84074. .107174
/note="assembly_fragment"
107275. .137395
/note="assembly_fragment"
137496. .146312
/note="assembly_fragment"
clone_end:r7
vector_side:right
BASE COUNT 37592 a 36284 c 35592 g 35732 t 1112 others
ORIGIN
Query Match 1.8%; Score 58; DB 2: Length 146312;
Best Local Similarity 100.0%; Pred. No. 2.1e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1187 ctgccaccttgcctcccaaatgctgggattacagccatgagccactgcgcaccagc 1244
|||||
Db 132968 CTGCCACCTTGGCCCTCCCAAAAGTCTGGGATTACAGGCATGAGCCACTGCGCCACG 133025
|||||
RESULT 44
AC015495
LOCUS
DEFINITION Homo sapiens clone RP11-21E14, WORKING DRAFT SEQUENCE, 13 unordered
pieces.
ACCESSION AC015495
VERSION AC015495.4 GI:10047765
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```


REFERENCE
AUTHORS

2 (bases 1 to 157599)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
 Chospel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 DeRellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
 O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
 Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
 Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
 and Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (25-DEC-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Dec 26, 2001 this sequence version replaced gi:14010824.
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5451

Center clone name: 233_C_13

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 30289: contig of 30289 bp in length
 * 30290 30389: gap of 100 bp
 * 30390 41503: contig of 11114 bp in length
 * 41504 41603: gap of 100 bp
 * 41604 107568: contig of 65965 bp in length
 * 107569 107668: gap of 100 bp
 * 107669 119412: contig of 11744 bp in length
 * 119413 119512: gap of 100 bp
 * 119513 139173: contig of 19661 bp in length
 * 139174 139273: gap of 100 bp
 * 139274 149978: contig of 10705 bp in length
 * 149979 150078: gap of 100 bp
 * 150079 157599: contig of 7521 bp in length.

FEATURES

Source

1. L57599
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="15"
 /clone="RP11-233C13"
 /clone_lib="RPC1-11 Human Male BAC"
 BASE COUNT 41693 a 38435 c 38160 g 38669 t 642 others
 ORIGIN

Query Match

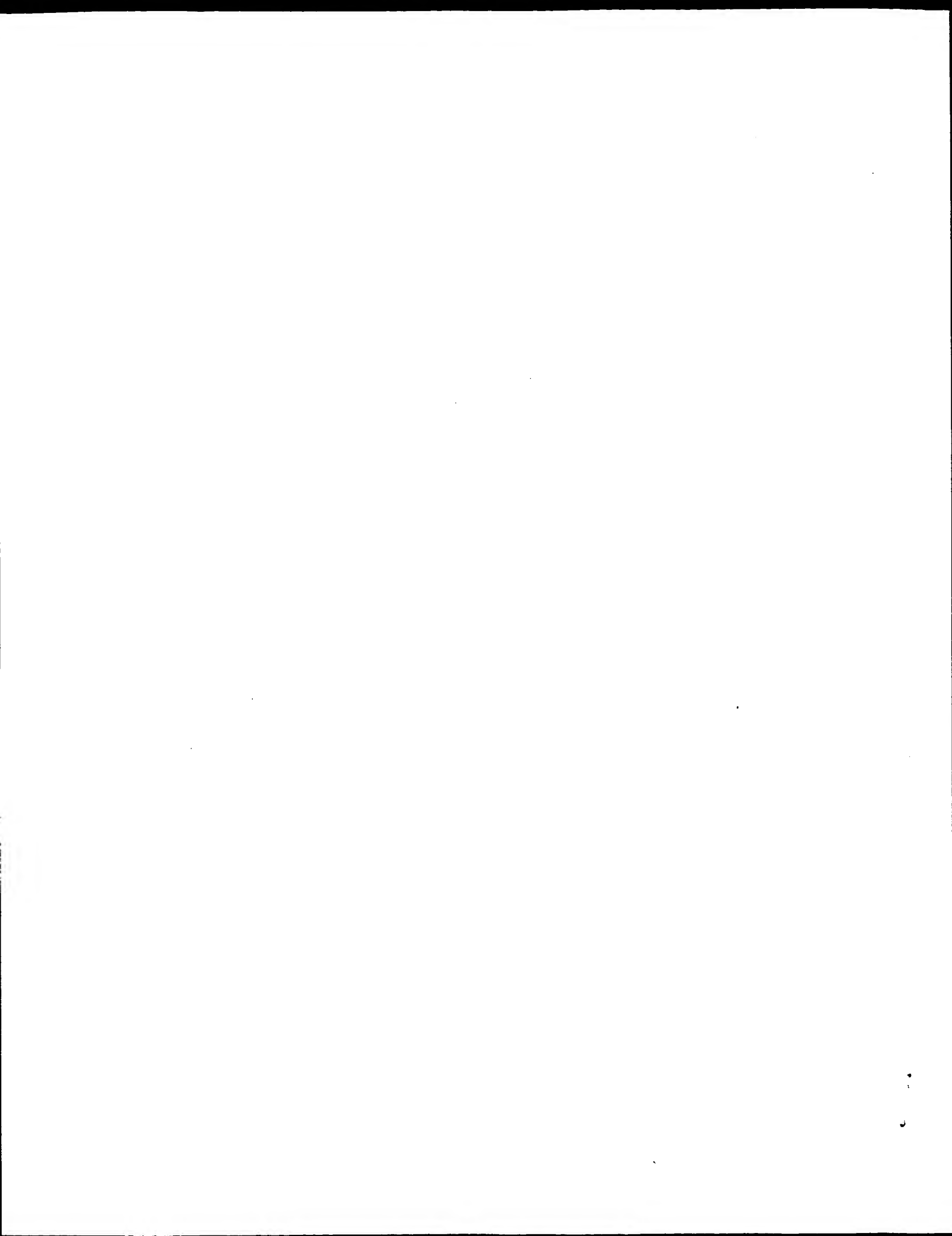
Best Local Similarity 1.8%; Score 58; DB 2; Length 157599;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggcctcccaagtctggattacaggcatgagccactgcccagc 1244

|||||

Db 145556 CTGCCACACTTTGGCCTCCCAAAGTGTGGGATTACAGGCATGAGCCACTGGGCCAGC 145613

Search completed: September 20, 2002, 06:25:43
 Job time: 18137 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:15:16 ; Search time 521.76 Seconds
(without alignments)
10631.997 Million cell updates/sec

Title: US-09-846-456-1
Perfect score: 3231
Sequence: 1 acagggcatgtgtgcagggtg.....gccccacatccccaccactt 3231

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	814	25.2	183999	22 AAF92831	Human ABC1 genomic
2	541	16.7	1643	22 AAF24681	Nucleotide sequenc
3	541	16.7	1643	22 AAF24703	Nucleotide sequenc
4	336	10.4	763	22 AAH04729	Human cDNA clone (
5	336	10.4	1750	22 AAH17451	Human cDNA sequenc
6	228	7.1	7260	22 AAD21326	Human ATP binding
7	228	7.1	7260	22 AAD170315	Human ATP binding
8	217	6.7	736	22 AAH07432	Human cDNA clone (
9	217	6.7	1556	22 AAH18606	Human cDNA sequenc

10	210	6.5	227	21 AAC09615	Human secreted pro
11	205	6.3	7086	22 ABA09200	Human ABCA1 homolo
12	205	6.3	7086	22 AAK52667	Human polynucleoti
13	205	6.3	7281	22 AAK51683	Human polynucleoti
14	205	6.3	9854	22 AAS06121	Human ABC1 DNA seq
15	197	6.1	10442	22 AAF24680	Nucleotide sequenc
16	197	6.1	10442	22 AAF24702	Nucleotide sequenc
17	188	5.8	10474	22 AAF24685	Nucleotide sequenc
18	188	5.8	10474	22 AAF24686	Nucleotide sequenc
19	188	5.8	10474	22 AAF24707	Nucleotide sequenc
20	188	5.8	10474	22 AAF24708	Nucleotide sequenc
21	92	2.8	446	22 AAS04035	Partial human ABC1
22	92	2.8	9741	22 AAS06120	Human ABC1 DNA seq
23	55	1.7	235033	19 AAV57926	Hereditary haemoch
24	55	1.7	237326	19 AAV57903	Hereditary haemoch
25	52	1.6	1316	22 AAI64613	Human ribosomal S1
26	52	1.6	5351	22 AAS31466	Human DNA for a no
27	52	1.6	6461	22 AAS31467	Human immune/haema
28	52	1.6	8319	22 AAK65197	Human immune/haema
29	52	1.6	10901	22 AAL03236	Human reproductiv
30	52	1.6	29329	22 ABA18026	Human nervous syst
31	52	1.6	29329	22 ABA20511	Human immune/haema
32	52	1.6	29329	22 AAK70791	Human immune/haema
33	52	1.6	29329	22 AAK78512	Human secreted pro
34	51	1.6	149	21 AAC25949	Human polynucleoti
35	51	1.6	396	22 AAI81653	Human polynucleoti
36	51	1.6	425	22 AAI87958	Human immune/haema
37	51	1.6	1856	22 AAK77921	Human polypeptide-
38	51	1.6	2046	22 AAI64545	Human immune/haema
39	51	1.6	7759	22 AAK77916	Human ovarian and
40	51	1.6	17904	22 ABA07913	Human reproductive
41	51	1.6	17904	22 AAL03730	Human musculooskele
42	51	1.6	26591	22 AAL36313	Human DNA for a no
43	51	1.6	32186	22 AAS34422	Human thioredoxin
44	51	1.6	66566	21 AAS33450	Human histone deac
45	51	1.6	122186	22 AAC89560	

ALIGNMENTS

RESULT 1

AAF92831
ID AAF92831 standard; DNA; 183999 BP.
XX
AC AAF92831;
XX
DT 17-MAY-2001 (first entry)
DE Human ABC1 genomic DNA.
XX
KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS Homo sapiens.
XX
PN WO200115676-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-IB01492.
XX
PR 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
XX (XENO-) XENON GENETICS INC.
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX WPI; 2001-244356/25.
DR Treating a lower than normal high density lipoprotein-cholesterol
XX
PT

PT (HDL-C) level, a higher than normal triglyceride level, or a cardiovascular disease, by administering a compound that modulates LXR-PT or RXR-mediated transcriptional activity -

PS Claim 8; Fig 1; 317pp; English.

The present invention relates to a method for treating a patient diagnosed as having a lower than normal high density lipoprotein-cholesterol (HDL-C) level, a higher than normal triglyceride level, or a cardiovascular disease, involving administering a compound that modulates LXR- or RXR-mediated transcriptional activity or ABC1 expression or activity. The LXR gene product may be used in an assay to identify compounds useful for the treatment of a disease or condition selected a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 25.2%; Score 814; DB 22; Length 183999;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 964; Conservative 0; Mismatches 3; Indels 0; Gaps

QY	782	atgaacacgagcgacgaagacatttaoataattgctcatcatgtgtgtgtoacagtttgac	841
Db	26671	atgaacacgagcgacgaagacatttaoataattgctcatcatgtgtgtgtoacagtttgac	
QY	842	cccaaaacccaattattgacccaaggtattcttttctgacgagccaaggggtccgcctctc	901
Db	26731	cccaaaacccaattattgacccaaggtattcttttctgacgagccaaggggtccgcctctc	26790
QY	902	ctgggacctgggcttttagaagctcatctctgacctcttctgagatccatcccttctcttt	961
Db	26791	ctgggacctgggcttttagaagctcatctctgacctcttctgagatccatcccttctcttt	26850
QY	962	tattttcttgacacgagcttgctctgacctcaggctggagtgagtgagtgagtgatct	1021
Db	26851	tattttcttgacacgagcttgctctgacctcaggctggagtgagtgagtgagtgatct	26910
QY	1022	cgactcaactgaacctgcctccggggttcaagcgattctcctgcctcagcctctcgag	1081
Db	26911	cgactcaactgaacctgcctccggggttcaagcgattctcctgcctcagcctctcgag	26970
QY	1082	ataacaggcgccgcacaccacatctgactaaatttttgatttttagtaaaacacgggttt	1141
Db	26971	ataacaggcgctcgcaacacatctgactaaatttttgatttttagtaaaacacgggttt	27030
QY	1142	catcatgtggcgaagttggttttcgaactctgtaoctgaggtgagctgccacacttgccc	1201
Db	27031	catcatgtggcgaagttggttttcgaactctgtaoctgaggtgagctgccacacttgccc	27090
QY	1202	tcccaaatgctggtggattacaggcatgagccactgcgccagctcagatccatcccttc	1261
Db	27091	tcccaaatgctggtgattacaggcatgagccactgcgccagctcagatccatcccttc	27150
QY	1262	taaggggcaaacagtcocatggtgcgaagggggccatgccacccagagttatgagtacctggg	1321
Db	27151	taaggggcaaacagtcocatggtgcgaagggggccatgccacccagagttatgagtacctggg	27210
QY	1322	actccagaattccttgctggtggcctccacatgcaactccagggcctgctgggcctct	1381
Db	27211	actccagaattccttgctggtggcctccacatgcaactccagggcctgctgggcctct	27270
QY	1382	tctatgcgtctgctgagttgttagaaccaactgatgtagtaacctgggcttgagccg	1441
Db	27271	tctatgggtctgctgagttgttagaaccaactgatgtagtaacctgggcttgagccg	27330
QY	1442	tggcctggagatcctgtgaactgtagcatgagggggcttgtgacgtgaatgtctgcac	1501
Db	27331	tggcctggagatcctgtgaactgtagcatgagggggcttgtgacgtgaatgtctgcac	27390
QY	1502	gcaggtgggtgggagttcttggaaatatgatgtgagctgaggtgggaagagagtaggctgg	1561

Db	27391	g c a g g t g g t g g g a g t t c t g g a a t a t g a t g g a g c t g a g g t g g g a a g a g a g t a g g c t t g g	27450
QY	1562	g g c a g c t c t c a t g c c a c c t c a t t c t g g c a a a c t c a g a g t c a a c t g t g a a g a g t a	1621
Db	27451	g g c a g c t c t c a t g c c a c c t c a t t c t g g c a a a a c t c a g g t c a a a c t g t g a a g a g t a	27510
QY	1622	a a t g t a a c t g c c c t t c a a g t g g g t c a a a g t a t c t t g c a a g t a g g a g a c c t t g	1681
Db	27511	a a t g t a a c t g c c c t t c a a g t g g g t c a a a g t a t c t t g c a a g t a g g a c c t t g	27570
QY	1682	t g g c t c a a c g t g a c t t c c a g g g c t g t g g c c t t c t a c g g g t c t c t c c t g a g t c	1741
Db	27571	t g g c t c a a c g t g a c t t c c a g g g c t g t g g c c t t c t a c g g g t c t c t c c t g a g t c	27630
QY	1742	t t c t a t g	1748
Db	27631	t t c t a t g	27637

RESULT 2
AAF24681
ID AAF24681 standard; DNA: 1643 BP.

AA
AC AAF24681;

XX
DT 20-APR-2001 (first entry)

DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

Human; adenosine triphosphate binding cassette protein 1; ABC1;
apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
chromosome 9q22-9q31; heart disease; hypercholesterolemia;
atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX
PN WO200078972-A2.XX
PD 28-DEC-2000.XX
PF 16-JUN-2000; 2000WO-US16765.XX
PR 18-JUN-1999; 99US-0140264.

PR	14-SEP-1999;	99US-0153872.
PR	19-NOV-1999;	99US-0166573.

XX
PA (CVTH-) CV THERAPEUTICS INC.XX
PI Lawn RM, Wade D, Garvin M:XX
DR
WPI: 2001-137812/14.

xx
PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT useful for the development of agents for the treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
PT

XX
PS
Claim 1: Page 143-144: 215pp: English.

The present sequence represents the 5' flanking region of the human adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1 resides in cell membranes and utilises ATP hydrolysis to transport a wide variety of substrates across the plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated mobilisation of intracellular cholesterol stores. ABC1 is defective in Tangier disease, a genetic disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome 9q22-q31. The ABC1 genes and proteins are useful for developing pharmaceutical agents for the treatment of heart disease and other disorders associated with hypercholesterolemia and atherosclerosis. The genes are useful for developing screening assays to screen for compounds that regulate the expression of genes associated with cholesterol transport. The genes and proteins are also useful for

Qy 2610 gaagggagcagaccgagaccctaagacacacctgtgtaccctccaccaccacc 2669
 Db 1274 gaagggagcagaccgagaccctaagacacacctgtgtaccctccaccaccacc 1333
 Qy 2670 c 2670
 Db 1334 c 1334

RESULT 4
 AAH04729
 ID AAH04729 standard; cDNA; 763 BP.
 AC AAH04729;
 XX
 XX 26-JUN-2001 (first entry)
 DT
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:1564.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 1; SEQ ID 1564; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

SQ Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

Query Match 10.4%; Score 336; DB 22; Length 763;
 Best Local Similarity 100.0%; Pred. No. 4.1e-143;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2896 aattgcgagcagagtgagtgaggcgagcccgagagccgagccgaccttctctcc 2955
 Db 1 aattgcgagcagagtgagtgaggcgagcccgagagccgagccgaccttctctcc 60
 Qy 2956 gggctgcggcagggcagggcgaggctccgcgaccacagagccggtctcaggcgcc 3015
 Db 61 gggctgcggcagggcagggcgaggctccgcgaccacagagccggtctcaggcgcc 120
 Qy 3016 ttgtctctgttttttcccggttctgttttctcccttcccttccggaaggtctcagg 3075
 Db 121 ttgtctctgttttttcccggttctgttttctcccttcccttccggaaggtctcagg 180
 Qy 3076 ggtagagaaaagagcgcgcaaacacacaaagtggaaacaggttaagaggtctctcagt 3135
 Db 181 ggtagagaaaagagcgcgcaaacacacaaagtggaaacaggttaagaggtctctcagt 240
 Qy 3135 tacttgggcggtattgttttctcgagggcgaagagcgttcggaaggtctcggttctcg 3195
 Db 241 tacttgggcggtattgttttctcgagggcgaagagcgttcggaaggtctcggttctcg 300
 Qy 3196 gggactttgatccgagccacacatccccaccatt 3231
 Db 301 gggactttgatccgagccacacatccccaccatt 336

RESULT 5

AAH17451
 ID AAH17451 standard; cDNA; 1750 BP.

XX AC AAH17451;
 XX

XX 26-JUN-2001 (first entry)
 DT

XX Human cDNA sequence SEQ ID NO:16905.
 DE

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW

XX Homo sapiens.
 OS

XX EP1074617-A2.
 PN

XX 07-FEB-2001.
 PD

XX 28-JUL-2000; 2000EP-0116126.
 PF

XX 29-JUL-1999; 99JP-0248036.
 PR

XX 27-AUG-1999; 99JP-0300253.
 PR

XX 11-JAN-2000; 2000JP-0118776.
 PR

XX 02-MAY-2000; 2000JP-0183767.
 PR

XX 09-JUN-2000; 2000JP-0241899.
 PR

XX (HELI-) HELIX RES INST.
 PA

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI

XX WPI; 2001-318749/34.
 DR

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 8; SEQ ID 16905; 2537pp + CD ROM; English.
 PS

XX The present invention describes primer sets for synthesising 5602
 CC

CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

SQ Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;

Query Match 10.4%; Score 336; DB 22; Length 1750;
 Best Local Similarity 100.0%; Pred. No. 4,1e-143; Indels 0; Gaps 0;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2896 aattgcagcagagtagtgaggccgagccgagccgagccgagccgagcccttctccc 2955
 Db 1 aattgcagcagagtagtgaggccgagccgagccgagccgagccgagcccttctccc 60
 QY 2956 gggctgcggcaggcaggcgggagctccgcaccacacagagccggttcaggggc 3015
 Db 61 gggctgcggcaggcaggcgggagctccgcaccacacagagccggttcaggggc 120
 QY 3016 ttgtctctgtttttcccggttctgttttcccttctccgaaagcttgcagg 3075
 Db 121 ttgtctctgtttttcccggttctgttttcccttctccgaaagcttgcagg 180
 QY 3076 ggtaggagaagagacacacacaaagtggaaacaggttaagagctctccagtga 3135
 Db 181 ggtaggagaagagacacacacaaagtggaaacaggttaagagctctccagtga 240
 QY 3136 tacttggcgttatgtttgttcgagcacaaggaggttcgggaagtgcgtggttcg 3195
 Db 241 tacttggcgttatgtttgttcgagcacaaggaggttcgggaagtgcgtggttcg 300
 QY 3196 gggactttgatccggagcccaatcccccactt 3231
 Db 301 gggactttgatccggagcccaatcccccactt 336

RESULT 6

AAAD21326
 ID AAD21326 standard; DNA; 7260 BP.

XX AC AAD21326;

XX 28-JAN-2002 (first entry)

XX Human ATP binding cassette transporter 1 (ABCI) gene.

XX Human; ATP binding cassette transporter 1; ABCI; coronary heart disease;
 KW dermatological; atherosclerosis; cardiovascular; inflammatory disease;
 KW psoriasis; lipid disorder; antibacterial; septic shock; gene therapy;
 KW immunosuppressive; lupus erythematosus; rheumatoid arthritis; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 321..7106
 FT CDS

FT /*tag= a
 FT /product= "Human ABCI protein"

XX EP1136552-A1.

XX 26-SEP-2001.

XX 20-MAR-2000; 2000EP-0105820.

XX 20-MAR-2000; 2000EP-0105820.

XX (FARB) BAYER AG.

XX Schmitz G, Bodzioch M;

PI WPI: 2001-640388/74.

DR P-PSDB; AAE13022.

XX New adenosine triphosphate binding cassette transporter-1 gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 PT cardiovascular diseases and inflammatory diseases

PS Example 1; Fig 1; 48pp; English.

XX The invention relates to four common polymorphisms in the gene encoding
 CC ATP-binding cassette transporter-1 (ABCI). ABCI is associated with
 CC decreased ApoA-1 mediated efflux of cholesterol. The polymorphisms in
 CC ABCI directly affects cellular lipid homeostasis, which is a key factor
 CC in the atherogenic processes. The ABCI polymorphisms are useful for
 CC diagnosing and treating lipid disorders, cardiovascular diseases
 CC (coronary heart disease, atherosclerosis) and inflammatory diseases
 CC (psoriasis, lupus erythematosus). The identification of ABCI as a
 CC transporter for interleukin-beta (IL-beta) identifies this gene as
 CC a candidate for treatment of inflammatory diseases including rheumatoid
 CC arthritis and septic shock. The present sequence is human ABCI gene.

SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 7.1%; Score 228; DB 22; Length 7260;
 Best Local Similarity 100.0%; Pred. No. 7.7e-94;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2887 aaaccccgtaattgcagcagagtagtgaggccgagccgagccgagccgagccccc 2946

Db 1 aaaccccgtaattgcagcagagtagtgaggccgagccgagccgagccgagccccc 60

QY 2947 ttctctcccggttcggtgcggcaggcaggcgggagctccgcgacacacagagccggttc 3006

Db 61 ttctctcccggttcggtgcggcaggcaggcgggagctccgcgacacacagagccggttc 120

QY 3007 tcaggcgtttgctctctgtttttcccggttctgttttcccttctccggaagtc 3066

Db 121 tcaggcgtttgctctctgtttttcccggttctgttttcccttctccggaagtc 180

QY 3067 ttgtcaagggttaggagaaagagacgacacacaaagtggaaaacag 3114

Db 181 ttgtcaagggttaggagaaagagacgacacacaaagtggaaaacag 228

RESULT 7

AAI70315

ID AAI70315 standard; cDNA; 7260 BP.

XX AC AAI70315;

XX 07-JAN-2002 (first entry)

XX Human ATP binding cassette transporter 1 (ABCI) cDNA.

XX ATP binding cassette transporter 1; ABCI; human; lipid disorder;
 KW cholesterol; cardiovascular disease; inflammatory disease;
 KW antiinflammatory; antilipemic; antipsoriatic; dermatological;

KW Tangier disease; coronary heart disease; diagnosis; gene therapy;
 XX polymorphism; ss.
 OS Homo sapiens.
 FH Location/Qualifiers
 FT CDS 321..7106
 FT CDS /*tag= a
 FT CDS 501..7106
 FT CDS /*tag= b
 FT variation /*note= "alternative open reading frame of AAI70314"
 FT variation replace(976,A)
 FT variation /*tag= c
 FT variation replace(1516,C)
 FT variation /*tag= d
 FT variation replace(2969,G)
 FT variation /*tag= e
 FT variation replace(3836,C)
 FT variation /*tag= f
 XX
 PN EP1136554-A1.
 PD 26-SEP-2001.
 XX
 PF 24-MAR-2000; 2000EP-0106401.
 XX
 PR 24-MAR-2000; 2000EP-0106401.
 XX
 XX (FARB) BAYER AG.
 XX
 XX Schmitz G, Bodzioch M;
 XX
 XX WPI; 2001-640389/74.
 DR P-PSDB; AAM50228.
 XX
 PT New adenosine triphosphate binding cassette transporter gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 PT cardiovascular diseases and inflammatory diseases
 XX
 PS Disclosure; Page 26-28; 41pp; English.
 XX
 CC The present sequence is that of cDNA encoding the human adenosine
 CC triphosphate (ATP) binding cassette transporter 1 (ABCI) protein
 CC (see AAM50227). The sequence includes an extended open reading
 CC frame (ORF) to that provided by the sequence in AAI70314, using
 CC an alternative ATG codon as initiation codon and thereby adding an
 CC extra 40 N-terminal amino acids to the encoded ABCI protein (see
 CC AAM50228). The invention provides 4 common polymorphisms in the
 CC ABCI gene. These were identified by sequencing the ABCI gene in
 CC different Tangier kindreds. In the variant genes (numbering as in
 CC AAI70314), G is changed to A at position 596, T is changed to C at
 CC position 1136, A is changed to G at position 2589 or G is changed
 CC to C at position 3456, or any combination of these. All of these
 CC polymorphisms alter the amino acid sequence of ABCI and therefore
 CC may affect its function. The 2 most common polymorphisms (G596A)
 CC and A2589G) are both associated with a decreased in vitro ApoA-I
 CC mediated efflux of cholesterol from mononuclear phagocytes, a
 CC feature typical of Tangier disease. 3 Of the variants (G596A,
 CC A2589G and G3456C) are significantly increased in a population of
 CC men having low high density lipoprotein-cholesterol levels and
 CC established coronary heart disease (CHD) relative to CHD-free
 CC control subjects. The use of the provided ABCI polymorphisms for
 CC the diagnosis and treatment of lipid disorders, cardiovascular
 CC diseases, and inflammatory diseases (e.g. psoriasis, lupus
 CC erythematoses) is claimed. Modulation of ABCI transcripts or
 CC proteins by antisense or ribozyme technology or RNA decoys is also
 CC claimed.
 XX
 SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 7.18; Score 228; DB 22; Length 7260;
 Best Local Similarity 100.0%; Pred. No. 7.7e-94;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2887 aaaccccgtaattgcgagcgagagtgagtgaggccgagccgagagccgagccgaccc 2946
 Db 1 aaaccccgtaattgcgagcgagagtgagtgaggccgagccgagagccgagccgaccc 60
 QY 2947 ttctctcccggtgctcgaggcagggcagggcgggagctccgagcaccacagagccggttc 3006
 Db 61 ttctctcccggtgctcgaggcagggcagggcgggagctccgagcaccacagagccggttc 120
 QY 3007 tcaggcgcttcgggaagc 3066
 Db 121 tcaggcgcttcgggaagc 180
 QY 3067 ttgtcaagggttaggagagagacgcaaacacaaagtggaacacag 3114
 Db 181 ttgtcaagggttaggagagagacgcaaacacaaagtggaacacag 228
 RESULT 8
 AAH07432
 ID AAH07432 standard; cDNA; 736 BP.
 XX
 AC AAH07432;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:4267.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 OS
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 6.7%; Score 217; DB 22; Length 736;
 Best Local Similarity 100.0%; Pred. No. 8.1e-89;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2898 ttgcagcagagtgagtgaggccgggacccgagagccgagccgaccccttctcccg 2957
 Db 5 ttgcagcagagtgagtgaggccgggacccgagagccgagccgaccccttctcccg 64
 QY 2958 gctgcgagcagggcagggcgagtcgcgcacacagagccggttctcagggcgctt 3017
 Db 65 gctgcgagcagggcagggcgagtcgcgcacacagagccggttctcagggcgctt 124
 QY 3018 tgctcctgtttttcccggttcttcttctccttctcggaaaggttgcagggg 3077
 Db 125 tgctcctgtttttcccggttcttcttctccttctcggaaaggttgcagggg 184
 QY 3078 taggagaaagagcagcaacacacaaagtggaaacag 3114
 Db 185 taggagaaagagcagcaacacacaaagtggaaacag 221

RESULT 9
 AAH18606
 ID AAH18606 standard; cDNA; 1556 BP.
 XX
 AC AAH18606;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:18808.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.
 XX

CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

Query Match 6.7%; Score 217; DB 22; Length 1556;
 Best Local Similarity 100.0%; Pred. No. 8.1e-89;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2898 ttgcagcagagtgagtgaggccgggacccgagagccgagccgaccccttctcccg 2957
 Db 5 ttgcagcagagtgagtgaggccgggacccgagagccgagccgaccccttctcccg 64
 QY 2958 gctgcgagcagggcagggcgagtcgcgcacacagagccggttctcagggcgctt 3017
 Db 65 gctgcgagcagggcagggcgagtcgcgcacacagagccggttctcagggcgctt 124
 QY 3018 tgctcctgtttttcccggttcttcttctccttctcggaaaggttgcagggg 3077
 Db 125 tgctcctgtttttcccggttcttcttctccttctcggaaaggttgcagggg 184
 QY 3078 taggagaaagagcagcaacacacaaagtggaaacag 3114
 Db 185 taggagaaagagcagcaacacacaaagtggaaacag 221

RESULT 10
 AAC09615
 ID AAC09615 standard; cDNA; 227 BP.
 XX
 AC AAC09615;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 13690.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;

```

XX WPI: 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1: SEQ ID 13690; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
XX Sequence 227 BP; 44 A; 65 C; 73 G; 45 T; 0 other;

Query Match      6.5%; Score 210; DB 21; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.3e-85;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2881 cggcaaaaccccaattgcagcagagtgagtggcgccgagccgcagagcgcgc 2940
Db 18 cggcaaaaccccaattgcagcagagtgagtggcgccgagccgcagagcgcgc 77
QY 2941 cgaccttctcccggtctgcgcagcgagcgagcgagcgagcgagcgagc 3000
Db 78 cgaccttctcccggtctgcgcagcgagcgagcgagcgagcgagcgagc 137
QY 3001 cggttctcaggcgcttgcctctgtttttcccggttctgttttcccttctccg 3060
Db 138 cggttctcaggcgcttgcctctgtttttcccggttctgttttcccttctccg 197
QY 3061 gaagcgttgcagggtgaggaagaga 3090
Db 198 gaagcgttgcagggtgaggaagaga 227

RESULT 11
ABA09200
ID ABA09200 standard; cDNA; 7086 BP.
AC ABA09200;
XX
XX 11-JAN-2002 (first entry)
XX
XX Human ABCA1 homologue-encoding cDNA, SEQ ID NO:976.
XX
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
XX cytotstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
XX antifungal; vulnuary; antiulcer; ss.
XX
XX Homo sapiens.

```

```

XX WO200157188-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-457740/49.
XX P-PSDB; AB111956.
XX
XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
XX
XX Claim 1: Page 833-835; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides,
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis; cancer cell proliferation or metastasis.
XX Depending on their biological activities, polypeptides and nucleotides of
XX the invention are useful for preventing, treating or ameliorating medical
XX conditions, e.g., by protein or gene therapy. Such conditions include
XX cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
XX proliferative retinopathy, atherosclerosis, coronary heart disease,
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
XX vascular growth. Polypeptides involved with tissue regeneration and
XX repair (or nucleic acids encoding them) may be used to promote wound
XX healing (e.g., of burns, incisions and ulcers), while those with
XX immunomodulatory activities may be used in the treatment of viral,
XX bacterial and fungal infections in addition to immune disorders.
XX Polypeptides with growth factor activity may be used in cell cultures to
XX promote cell growth. For example, such polypeptides may be used to
XX manipulate stem cells in culture to give rise to neuroepithelial cells
XX that can be used to augment or replace cells damaged by illness,
XX autoimmune disease or accidental damage. The polypeptides and nucleotides
XX may also be used in the diagnosis of the above conditions, and in drug
XX screening techniques. The present sequence represents a cDNA encoding a
XX novel human polypeptide of the invention.
XX
XX Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match      6.3%; Score 205; DB 22; Length 7086;
Best Local Similarity 100.0%; Pred. No. 2.4e-83;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtggcgccgagcccgagagccgagccgagcccttctcccggtcgcgagg 2969
Db 7 gtgagtggcgccgagcccgagagccgagcccttctcccggtcgcgagg 66
QY 2970 cagggcggggagctccgcgcacacacagagccggttctcgcgcttctctctt 3029

```


CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 6.3%; Score 205; DB 22; Length 7281;
 Best Local Similarity 100.0%; Pred. No. 2.4e-83;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtggcgccggagcccgagagccgagccgagccctctctcccggtcgcgaggg 2969
 |||||||
 Db 45 gtgagtggcgccggagcccgagagccgagccgagccctctctcccggtcgcgaggg 104
 QY 2970 caggcggggagctccgcgacacagagccggttctcaggcgctttgctctgttt 3029
 |||||||
 Db 105 caggcggggagctccgcgacacagagccggttctcaggcgctttgctctgttt 164
 QY 3030 ttcccggttctgtttctccctctccggaagctgtcgaaggtaggagaaagag 3089
 |||||||
 Db 165 ttcccggttctgtttctccctctccggaagctgtcgaaggtaggagaaagag 224
 QY 3090 acgcaaacacaaaagtggaaaacag 3114
 |||||||
 Db 225 acgcaaacacaaaagtggaaaacag 249

RESULT 14

AA06121
 ID AAS06121 standard; cDNA; 9854 BP.

AC AAS06121;

DT 12-SEP-2001 (first entry)

DE Human ABC1 DNA sequence #2.

XX Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
 KW cardiovascular; neurological; Tangier disease; LCAT deficiency;
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 298..7078
 FT /*tag= a
 FT /product= "Human ABC1 protein"

PN WO200130848-A2.

XX 03-MAY-2001.

XX 26-OCT-2000; 2000WO-EF10886.

XX 26-OCT-1999; 99EP-0402668.

PR 01-MAR-2000; 2000US-0186260.

XX (AVET) AVENTIS PHARMA SA.

PA Denefle P, Rosier-Montus M, Arnould-Requigne I, Prades C, Naudin L;

PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;

PI Dean M;

XX WPI; 2001-316327/33.

DR

DR P-PSDB; AAO02176.

XX New human ABC1 nucleic acids and polypeptides for treating
 PT atherosclerosis, malaria and diabetes -

XX Claim 1; Page 209-213; 368pp; English.

XX The sequence represents the coding sequence #2 of human ABC1. The
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,
 CC and polypeptides and vectors are useful for the prevention of
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is
 CC useful for screening for an active ingredient for the prevention or
 CC treatment of a disease resulting from dysfunction in the reverse
 CC transport of cholesterol. The nucleic acids and polypeptides are also
 CC useful for treating and preventing cardiovascular and neurological
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-
 CC cholesterol (LCAT) deficiency, malaria and diabetes.

XX Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;

Query Match 6.3%; Score 205; DB 22; Length 9854;
 Best Local Similarity 100.0%; Pred. No. 2.4e-83;

Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtggcgccggagcccgagagccgagccgagccctctctcccggtcgcgaggg 2969
 |||||||

Db 1 gtgagtggcgccggagcccgagagccgagccgagccctctctcccggtcgcgaggg 60
 |||||||

QY 2970 caggcggggagctccgcgacacagagccggttctcaggcgctttgctctgttt 3029
 |||||||

Db 61 caggcggggagctccgcgacacagagccggttctcaggcgctttgctctgttt 120
 |||||||

QY 3030 ttcccggttctgtttctccctctccggaagctgtcgaaggtaggagaaagag 3089
 |||||||

Db 121 ttcccggttctgtttctccctctccggaagctgtcgaaggtaggagaaagag 180
 |||||||

QY 3090 acgcaaacacaaaagtggaaaacag 3114
 |||||||

Db 181 acgcaaacacaaaagtggaaaacag 205
 |||||||

RESULT 15

AAF24680

ID AAF24680 standard; DNA; 10442 BP.

XX AAF24680;

XX 20-APR-2001 (first entry)

DE Nucleotide sequence of a human ABC1 polypeptide.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 291..7076
 FT /*tag= a
 FT /product= "ABC1 polypeptide"

PN WO200078972-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

PR

```
PR 19-NOV-1999; 99US-0166573.
XX (CVTH-) CV THERAPEUTICS INC.
XX PA
XX PI Lawn RM, Wade D, Garvin M;
XX DR WPI; 2001-137812/14.
XX XX
XX XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
XX PT useful for the development of agents for the treatment of heart disease
XX PT and other disorders associated with hypercholesterolemia and
XX PT atherosclerosis -
XX XX
XX PS Disclosure; Page 122-128; 215pp; English.
XX XX
XX CC The present sequence encodes a human adenosine triphosphate (ATP)
XX CC binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
XX CC membranes and utilises ATP hydrolysis to transport a wide variety of
XX CC substrates across the plasma membrane. ABC1 is a pivotal protein in
XX CC the apolipoprotein-mediated mobilisation of intracellular cholesterol
XX CC stores. ABC1 is defective in Tangier disease, a genetic disorder
XX CC characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
XX CC localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
XX CC useful for developing pharmaceutical agents for the treatment of heart
XX CC disease and other disorders associated with hypercholesterolemia and
XX CC atherosclerosis. The genes are useful for developing screening assays to
XX CC screen for compounds that regulate the expression of genes associated
XX CC with cholesterol transport. The genes and proteins are also useful for
XX CC other disorders associated with hypercholesterolemia.
XX XX
XX SQ Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 6.1%; Score 197; DB 22; Length 10442;
Best Local Similarity 100.0%; Pred. NO. 1.1e-79;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2918 ggcgggaccgcagagccgagccgacctctctccgggtcgccgagggcagggcg 2977
DB 1 ggcgggaccgcagagccgagccgacctctctccgggtcgccgagggcagggcg 60
QY 2978 ggaagtcgcgcacacacagacgggtctcaggcgcttctgtctctgtttttcccg 3037
DB 61 ggaagtcgcgcacacacagacgggtctcaggcgcttctgtctctgtttttcccg 120
QY 3038 gttctgttttctccctctccggaaggcttgcaggggtaggagaagacgcaaac 3097
DB 121 gttctgttttctccctctccggaaggcttgcaggggtaggagaagacgcaaac 180
QY 3098 acaaaagtggaaaacag 3114
DB 181 acaaaagtggaaaacag 197

RESULT 16
AAF24702
ID - AAF24702 standard; DNA; 10442 BP.
XX AC
XX AC AAF24702;
XX XX
XX DT 20-APR-2001 (first entry)
XX XX
XX DE Nucleotide sequence of a human ABC1 polypeptide.
XX KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
XX KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
XX KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
XX KW atherosclerosis; cholesterol transport; ss.
XX XX
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
```

```
FT CDS 291..7076
FT FT /*tag= a
XX XX /product= "ABC1 polypeptide"
XX PN WO200078971-A2.
XX PD 28-DEC-2000.
XX XX
XX PF 16-JUN-2000; 2000WO-US16591.
XX PR 18-JUN-1999; 99US-0140264.
XX PR 14-SEP-1999; 99US-0153872.
XX PR 19-NOV-1999; 99US-0166573.
XX XX
XX PA (CVTH-) CV THERAPEUTICS INC.
XX PA (UNIW ) UNIV WASHINGTON.
XX PI Lawn RM, Wade D, Oram JF, Garvin M;
XX XX
XX DR WPI; 2001-137811/14.
XX DR P-PSDB; AAB31365.
XX XX
XX PT Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX PT polynucleotides and polypeptides, useful for treatment of heart disease
XX PT and other disorders associated with hypercholesterolemia and
XX PT atherosclerosis -
XX XX
XX PS Claim 3; Page 117-123; 211pp; English.
XX XX
XX CC The present sequence encodes a human adenosine triphosphate (ATP)
XX CC binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
XX CC membranes and utilises ATP hydrolysis to transport a wide variety of
XX CC substrates across the plasma membrane. ABC1 is a pivotal protein in
XX CC the apolipoprotein-mediated mobilisation of intracellular cholesterol
XX CC stores. ABC1 is defective in Tangier disease, a genetic disorder
XX CC characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
XX CC localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
XX CC useful for developing pharmaceutical agents for the treatment of heart
XX CC disease and other disorders associated with hypercholesterolemia and
XX CC atherosclerosis. The genes are useful for developing screening assays to
XX CC screen for compounds that regulate the expression of genes associated
XX CC with cholesterol transport. The genes and proteins are also useful for
XX CC other disorders associated with hypercholesterolemia.
XX XX
XX SQ Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 6.1%; Score 197; DB 22; Length 10442;
Best Local Similarity 100.0%; Pred. No. 1.1e-79;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2918 ggcgggaccgcagagccgagccgacctctctccgggtcgccgagggcagggcg 2977
DB 1 ggcgggaccgcagagccgagccgacctctctccgggtcgccgagggcagggcg 60
QY 2978 ggaagtcgcgcacacacagacgggtctcaggcgcttctgtctctgtttttcccg 3037
DB 61 ggaagtcgcgcacacacagacgggtctcaggcgcttctgtctctgtttttcccg 120
QY 3038 gttctgttttctccctctccggaaggcttgcaggggtaggagaagacgcaaac 3097
DB 121 gttctgttttctccctctccggaaggcttgcaggggtaggagaagacgcaaac 180
QY 3098 acaaaagtggaaaacag 3114
DB 181 acaaaagtggaaaacag 197

RESULT 17
AAF24685
ID - AAF24685 standard; DNA; 10474 BP.
XX XX
```

AC AAF24685;
 XX 20-APR-2001 (first entry)
 XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;
 XX apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 323..7108
 FT /*tag= a
 FT /product= "defective ABC1 polypeptide"
 FT
 PN WO200078972-A2.
 XX
 XX 28-DEC-2000.
 XX
 XX 16-JUN-2000; 2000WO-US16765.
 XX
 XX 18-JUN-1999; 99US-0140264.
 PR 14-SEP-1999; 99US-0153872.
 PR 19-NOV-1999; 99US-0166573.
 XX
 XX (CVTH-) CV THERAPEUTICS INC.
 XX Lawn RM, Wade D, Garvin M;
 XX WPI; 2001-137812/14.
 XX
 XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
 PT useful for the development of agents for the treatment of heart disease
 PT and other disorders associated with hypercholesterolemia and
 PT atherosclerosis -
 XX
 PS Disclosure; Page 148-154; 215pp; English.
 XX
 CC The present sequence encodes a human adenosine triphosphate (ATP)
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
 CC ATP hydrolysis to transport a wide variety of substrates across the
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
 CC Tangier disease, a genetic disorder characterised by abnormal
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
 CC 9q22-9q31. The ABC1 genes and proteins are useful for developing
 CC pharmaceutical agents for the treatment of heart disease and other
 CC disorders associated with hypercholesterolemia and atherosclerosis. The
 CC genes are useful for developing screening assays to screen for compounds
 CC that regulate the expression of genes associated with cholesterol
 CC transport. The genes and proteins are also useful for are also useful
 CC as diagnostic indicators of cardiovascular disease and other disorders
 CC associated with hypercholesterolemia.
 XX
 SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match 5.8%; Score 188; DB 22; Length 10474;
 Best Local Similarity 100.0%; Pred. No. 1.4e-75;
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccgacctctctccgggctgcgcagcagcagcggagctccg 2986
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 42 ccgcagagccgagccgacctctctccgggctgcgcagcagcagcggagctccg 101
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2987 cgcacacacagcgggtctcagggcgcttgcctctgtttttcccggtctgttt 3046
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 102 cgcacacacagcgggtctcagggcgcttgcctctgtttttcccggtctgttt 161
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 3047 tctcccttctccggaagcgttgcacagggttagagagaagacgcaacacaaaagtg 3106
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 162 tctcccttctccggaagcgttgcacagggttagagagaagacgcaacacaaaagtg 221
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3107 gaaaacag 3114
 ||||||||
 Db 222 gaaaacag 229
 ||||||||
 RESULT 18
 AAF24686
 ID AAF24686 standard; DNA; 10474 BP.
 XX
 XX AC AAF24686;
 XX
 XX 20-APR-2001 (first entry)
 XX
 XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;
 XX apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.
 XX
 OS Homo sapiens.
 FH
 FH Key Location/Qualifiers
 FT 323..7108
 FT /*tag= a
 FT /product= "defective ABC1 polypeptide"
 FT
 XX WO200078972-A2.
 XX
 XX 28-DEC-2000.
 XX
 XX 16-JUN-2000; 2000WO-US16765.
 XX
 XX 18-JUN-1999; 99US-0140264.
 PR 14-SEP-1999; 99US-0153872.
 PR 19-NOV-1999; 99US-0166573.
 XX
 XX (CVTH-) CV THERAPEUTICS INC.
 XX Lawn RM, Wade D, Garvin M;
 XX WPI; 2001-137812/14.
 XX
 XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
 PT useful for the development of agents for the treatment of heart disease
 PT and other disorders associated with hypercholesterolemia and
 PT atherosclerosis -
 XX
 PS Disclosure; Page 170-176; 215pp; English.
 XX
 CC The present sequence encodes a human adenosine triphosphate (ATP)
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
 CC ATP hydrolysis to transport a wide variety of substrates across the
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
 CC Tangier disease, a genetic disorder characterised by abnormal
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
 CC 9q22-9q31. The ABC1 genes and proteins are useful for developing
 CC pharmaceutical agents for the treatment of heart disease and other
 CC disorders associated with hypercholesterolemia and atherosclerosis. The
 CC genes are useful for developing screening assays to screen for compounds
 CC that regulate the expression of genes associated with cholesterol
 CC transport. The genes and proteins are also useful for are also useful
 CC as diagnostic indicators of cardiovascular disease and other disorders
 CC associated with hypercholesterolemia.
 XX
 SQ Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

```
Query Match          5.8%; Score 188; DB 22; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.4e-75;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccacctctctccgggctgcgcagggcgagggcgagctccg 2986
DB 42 ccgcagagccgagccacctctctccgggctgcgcagggcgagggcgagctccg 101
QY 2987 cgcaccaacagagccggttctcaggcgcttgcctgttttttcccggtctgttt 3046
DB 102 cgcaccaacagagccggttctcaggcgcttgcctgttttttcccggtctgttt 161
QY 3047 tctcccttccggaagcgttgcaggggtaggaagagacgcaaacacaaaagt 3106
DB 162 tctcccttccggaagcgttgcaggggtaggaagagacgcaaacacaaaagt 221

QY 3107 gaaaacag 3114
DB 222 gaaaacag 229

RESULT 19
AAF24707
ID AAF24707 standard; DNA; 10474 BP.
AC AAF24707;
XX
DT 20-APR-2001 (first entry)
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 323..7108
FT FT /*tag= a
FT FT /product= "defective ABC1 polypeptide"
XX
PN WO200078971-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16591.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
PA (CVTH-) CV THERAPEUTICS INC.
PA (UNIW ) UNIV WASHINGTON.
XX
PI Lawn RM, Wade D, Oram JF, Garvin M;
XX
WPI: 2001-137811/14.
DR P-PSDB; AAB31366.
XX
PT Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT polynucleotides and polypeptides, useful for treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
PS Claim 27; Page 144-150; 21pp; English.
XX
CC The present sequence encodes a human adenosine triphosphate (ATP)
CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC ATP hydrolysis to transport a wide variety of substrates across the
```

```
CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC Tangier disease, a genetic disorder characterised by abnormal
CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC 9q22-9q31. The ABC1 genes and proteins are useful for developing
CC pharmaceutical agents for the treatment of heart disease and other
CC disorders associated with hypercholesterolemia and atherosclerosis. The
CC genes are useful for developing screening assays to screen for compounds
CC that regulate the expression of genes associated with cholesterol
CC transport. The genes and proteins are also useful for are also useful
CC as diagnostic indicators of cardiovascular disease and other disorders
CC associated with hypercholesterolemia.
XX
```

SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match 5.8%; Score 188; DB 22; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.4e-75;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2927 ccgcagagccgagccacctctctccgggctgcgcagggcgagggcgagctccg 2986
DB 42 ccgcagagccgagccacctctctccgggctgcgcagggcgagggcgagctccg 101
QY 2987 cgcaccaacagagccggttctcaggcgcttgcctgttttttcccggtctgttt 3046
DB 102 cgcaccaacagagccggttctcaggcgcttgcctgttttttcccggtctgttt 161
QY 3047 tctcccttccggaagcgttgcaggggtaggaagagacgcaaacacaaaagt 3106
DB 162 tctcccttccggaagcgttgcaggggtaggaagagacgcaaacacaaaagt 221
QY 3107 gaaaacag 3114
DB 222 gaaaacag 229
```

RESULT 20

```
AAF24708
ID AAF24708 standard; DNA; 10474 BP.
XX
AC AAF24708;
XX
DT 20-APR-2001 (first entry)
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 323..7108
FT FT /*tag= a
FT FT /product= "defective ABC1 polypeptide"
XX
PN WO200078971-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16591.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
PA (CVTH-) CV THERAPEUTICS INC.
PA (UNIW ) UNIV WASHINGTON.
XX
PI Lawn RM, Wade D, Oram JF, Garvin M;
```


PN WO200130848-A2.
 XX 03-MAY-2001.
 XX 26-OCT-2000; 2000WO-EP10886.
 XX 26-OCT-1999; 99EP-0402668.
 PR 01-MAR-2000; 2000US-0186260.
 XX (AVET) AVENTIS PHARMA SA.
 XX Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;
 PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;
 PI Dean M;
 XX WPI: 2001-316327/33.
 DR P-PSDB; RAU02176.
 XX New human ABC1 nucleic acids and polypeptides for treating
 PT atherosclerosis, malaria and diabetes -
 PT Claim 1; Page 204-208; 368pp; English.
 XX The sequence represents the coding sequence #1 of human ABC1. The
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,
 CC and polypeptides and vectors are useful for the prevention of
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is
 CC useful for screening for an active ingredient for the prevention or
 CC treatment of a disease resulting from dysfunction in the reverse
 CC transport of cholesterol. The nucleic acids and polypeptides are also
 CC useful for treating and preventing cardiovascular and neurological
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-
 CC cholesterol (LCAT) deficiency, malaria and diabetes.
 XX Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;
 SQ

Query Match 2.8%; Score 92; DB 22; Length 9741;
 Best Local Similarity 100.0%; Pred. No. 8.6e-32;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3023 cttgtttttcccggttctgtttttcccttcccggaaggtgttcaagggttagga 3082
 Db 1 cttgtttttcccggttctgtttttcccttcccggaaggtgttcaagggttagga 60
 Qy 3083 gaaagagacgcaaacacacaaagtgaacag 3114
 Db 61 gaaagagacgcaaacacacaaagtgaacag 92

RESULT 23
 AAV57926
 ID AAV57926 standard; DNA; 235033 BP.
 AC AAV57926;
 XX 23-DEC-1998 (first entry)
 XX Hereditary haemochromatosis subregion from an unaffected individual.
 XX Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
 KW diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
 KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
 KW type 1 sodium transport gene; ss.
 XX Homo sapiens.
 OS WO9814466-A1.
 PN 09-APR-1998.
 PD 30-SEP-1997; 97WO-US17658.
 PF

XX 07-MAY-1997; 97US-0852495.
 PR 01-OCT-1996; 96US-0724394.
 XX (PROG-) PROCENTIOR INC.
 XX Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
 PI Tsuchihashi Z, Wolff RK;
 XX WPI: 1998-240014/21.
 XX Hereditary haemochromatosis gene products - used to develop products
 PT for the diagnosis and treatment of hereditary disorders in iron
 PT metabolism
 XX Example 2; Fig 8; 209pp; English.
 XX The present invention describes hereditary haemochromatosis gene
 CC products from the human haemochromatosis gene. The present sequence
 CC represents a hereditary haemochromatosis subregion from an individual
 CC unaffected by hereditary haemochromatosis (HH). Also described is a
 CC method to determine the presence or absence of the common hereditary
 CC haemochromatosis (HFE) gene mutation in an individual comprising:
 CC (a) providing DNA or RNA from the individual; and (b) assessing the
 CC DNA or RNA for the presence or absence of a haplotype or genotype where
 CC the presence or absence of the haplotype genotype indicates the likely
 CC presence of the HFE gene mutation in the genome of the individual. The
 CC HFE gene sequences from the present invention can be used to develop
 CC products for use in the diagnosis and treatment of HFE. The present
 CC invention also describes BTF genes, which are homologues of the milk
 CC protein butyrophilin (BT), and can be used in the production of agonists
 CC and antagonists of BT function. Also described are: (1) a RoRet gene
 CC which can be used to develop products for the study, diagnosis and
 CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
 CC which are homologues of a type 1 sodium transport gene, and can
 CC similarly be used for hypophosphatemia.
 XX Sequence 235033 BP; 68786 A; 48466 C; 49441 G; 68340 T; 0 other;
 SQ

Query Match 1.7%; Score 55; DB 19; Length 235033;
 Best Local Similarity 100.0%; Pred. No. 6.4e-15;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1190 cccaccttgctcccaagtgctggattacagcatgagccactgcgccagc 1244
 Db 225009 cccaccttgctcccaagtgctggattacagcatgagccactgcgccagc 225063

RESULT 24
 AAV57903
 ID AAV57903 standard; DNA; 237326 BP.
 XX AAV57903;
 XX 21-DEC-1998 (first entry)
 XX Hereditary haemochromatosis subregion from an HH affected individual.
 XX Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
 KW diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
 KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
 KW type 1 sodium transport gene; ss.
 XX Homo sapiens.
 OS WO9814466-A1.
 PN 09-APR-1998.
 PD 30-SEP-1997; 97WO-US17658.
 PF 07-MAY-1997; 97US-0852495.
 PR

PR 01-OCT-1996; 9605-0724394.
 XX (PROG-) PROGENITOR INC.
 XX Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
 PI Tsuchihashi Z, Wolff RK;
 XX WPI; 1998-240014/21.
 DR Hereditary haemochromatosis gene products - used to develop products
 XX for the diagnosis and treatment of hereditary disorders in iron
 XX metabolism
 PS Claim 1; Fig 9; 209pp; English.
 XX The present invention describes hereditary haemochromatosis gene
 CC products from the human haemochromatosis gene. The present sequence
 CC represents a hereditary haemochromatosis subregion from an hereditary
 CC haemochromatosis (HH) affected individual. Also described is a
 CC method to determine the presence or absence of the common hereditary
 CC haemochromatosis (HFE) gene mutation in an individual comprising:
 CC (a) providing DNA or RNA from the individual; and (b) assessing the
 CC DNA or RNA for the presence or absence of a haplotype or genotype where
 CC the presence or absence of the haplotype genotype indicates the likely
 CC presence of the HFE gene mutation in the genome of the individual. The
 CC HFE gene sequences from the present invention can be used to develop
 CC products for use in the diagnosis and treatment of HFE. The present
 CC invention also describes BTF genes, which are homologues of the milk
 CC protein butyrophilin (BT), and can be used in the production of agonists
 CC and antagonists of BT function. Also described are: (1) a Roret gene
 CC which can be used to develop products for the study, diagnosis and
 CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
 CC which are homologues of a type 1 sodium transport gene, and can
 CC similarly be used for hypophosphatemia.
 XX
 SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 1.7%; Score 55; DB 19; Length 237326;
 Best Local Similarity 100.0%; Pred. No. 6.4e-15;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacttgccctcccaaaagtctggattacaggaatgagccactgcgccagc 1244
 |||||
 Db 227302 cccacttgccctcccaaaagtctggattacaggaatgagccactgcgccagc 227356

RESULT 25
 AAI64613
 ID AAI64613 standard; cDNA; 1316 BP.
 XX
 AC AAI64613;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Human ribosomal S11 protein l2 encoding cDNA.
 XX
 KW Human; ribosomal S11 protein l2; cytosolic; virucidal;
 KW immunomodulatory; antiinflammatory; haemostatic; malignant tumour;
 KW human immunodeficiency virus; HIV; infection; immunological disease;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 643..969
 FT /*tag= a
 FT /product= "ribosomal S11 protein l2"
 FT /note= "claimed in claim 6"
 XX
 PN W0200172801-A1.
 XX
 PD 04-OCT-2001.

XX 26-MAR-2001; 2001WO-CN00438.
 XX
 XX 27-MAR-2000; 2000CN-0115182.
 XX
 XX (SHAN-) SHANGHAI BLOWINDOW GENE DEV INC.
 XX
 XX Mao Y, Xie Y;
 XX
 XX WPI; 2001-597104/67.
 DR P-PSDB; AAG78169.
 DR
 XX New human ribosomal S11 protein l2 and encoded polynucleotide,
 PT applicable in diagnosis and treatment of malignant tumour, haemopathy,
 PT human immunodeficiency virus infection, immunological diseases and
 PT inflammation
 PT
 XX Claim 6; Page 29-30; 34pp; Chinese.
 PS
 XX The invention relates to the human ribosomal S11 protein l2 with
 CC cytosolic, virucidal, immunomodulatory, antiinflammatory and
 CC haemostatic activity. The protein and encoding polynucleotide are used
 CC in diagnosis and treatment of malignant tumour, haemopathy, human
 CC immunodeficiency virus (HIV) infection, immunological diseases and
 CC various inflammations. The polynucleotide is useful in gene therapy.
 XX
 SQ Sequence 1316 BP; 338 A; 306 C; 297 G; 375 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 1316;
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1193 accttggccctcccaaaagtctggattacaggaatgagccactgcgccagc 1244
 |||||
 Db 977 accttggccctcccaaaagtctggattacaggaatgagccactgcgccagc 1028

RESULT 26
 AAS31466/c
 ID AAS31466 standard; DNA; 5351 BP.
 XX
 AC AAS31466;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Human DNA for a novel extracellular matrix protein, Seq ID No 545.
 XX
 KW Human; secreted extracellular matrix protein; ds; immunomodulatory;
 KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
 KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
 KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
 KW Sezary syndrome; Gaucher's disease; neurological diseases;
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;
 KW wound healing; immunogen; gene therapy; antisense; food additive.
 XX
 OS Homo sapiens.
 XX
 PN W0200155368-A1.
 XX
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01348.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.

CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. For example, disorders associated with decreased
CC expression of SPs. The SP polynucleotide or a vector expressing them may
CC be administered to treat diseases by gene therapy. Antisense molecules
CC may be administered to down regulate expression of SPs by binding with
CC the cells own genes and preventing their expression. The polynucleotides
CC may also be used as DNA probes in diagnostic assays. The SPs may also be
CC used as antigens to produce antibodies and to identify modulators
CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC SP and as diagnostic agents for detecting the presence of SPs in samples.
CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
CC arrest, tachycardia and angina), infections caused by bacteria, viruses
CC and fungi and ocular disorders (e.g. corneal infections). Other uses
CC include wound healing, maintenance of organs before transplantation,
CC support of cell culture of primary tissues, modulation of for example

Query Match 1.68; Score 52; DB 22; Length 5351;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1193 accttgctcccaagtgctggattacaggcatgagccactgcgccagc 1244

|||||

Db 989 ACCITGGCTCCCAAGTGGGATTACAGGCATGAGCCACTGCGCCGACG 938

RESULT 27

AAS31467/C

ID AAS31467 standard; DNA; 6461 BP.

XX AC AAS31467;

XX DT 04-DEC-2001 (first entry)

XX DE Human DNA for a novel extracellular matrix protein, Seq ID No 546.

XX KW Human; secreted extracellular matrix protein; ds; immunomodulatory;
XX KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
XX KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
XX KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
XX KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
XX KW Sezary syndrome; Gaucher's disease; neurological diseases;
XX KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
XX KW cardiac arrest; tachycardia; angina; infection; corneal infections;
XX KW wound healing; immunogen; gene therapy; antisense; food additive.

XX OS Homo sapiens.

XX PN WO200155368-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01348.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465572/50.

XX Nucleic acid molecules encoding human secreted extracellular matrix
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
 PT Alzheimer's and Parkinson's diseases and cancers -

PS Claim 1; SEQ ID No 546; 577pp; English.

XX The invention relates to isolated nucleic acid molecules encoding
 CC novel human secreted extracellular matrix proteins (SPs). The
 CC polynucleotides and proteins are used to prevent, treat a medical
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 CC chickens or sheep. For example, disorders associated with decreased
 CC expression of SPs. The SP polynucleotide or a vector expressing them may
 CC be administered to treat diseases by gene therapy. Antisense molecules
 CC may be administered to down regulate expression of SPs by binding with
 CC the cells own genes and preventing their expression. The polynucleotides
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be

CC used as antigens to produce antibodies and to identify modulators
 CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
 CC antagonists may also be used to down regulate expression and activity of
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses
 CC include wound healing, maintenance of organs before transplantation,
 CC support of cell culture of primary tissues, modulation of for example

Query Match 1.6%; Score 52; DB 22; Length 6461;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1193 acctggcctcccaagtgctgggattacaggcatgagccactgcgccagc 1244
 |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 4277 ACCTTGGCCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGC 4226

RESULT 28

AAK65197

ID AAK65197 standard; DNA; 8319 BP.

XX AAK65197;

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20009.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytosstatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184564.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225477.

PR 14-AUG-2000; 2000US-0225575.

PR 14-AUG-2000; 2000US-02257558.
PR 14-AUG-2000; 2000US-02257559.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 23-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 20009; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.

Sequence 8319 BP; 2521 A; 1610 C; 1560 G; 2628 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 8319;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacttgccctcccaagtctgggattacagcgatgagccactggccc 1241
 |||||
Db 4718 cccacttgccctcccaagtctgggattacagcgatgagccactggccc 4769

RESULT 29

AAL03236

ID AAL03236 standard; DNA; 10901 BP.

XX AAL03236;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 5924.

KW Human; reproductive system related antigen; reproductive system disorder;

XX cancer; gene therapy; ds.

XX Homo sapiens.

PN WO200155320-A2.

PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01339.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.

PS Disclosure; SEQ ID NO 10357; 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (ABA11004-ABA21534) and proteins (AB114678-AB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 29329 BP; 5864 A; 6756 C; 7225 G; 8484 T; 0 other;

SQ

Query Match 1.6%; Score 52; DB 22; Length 29329;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattgctcccaagtgctgggattacagcagcactgcg 1238
|||||
DB 22210 CTGCCACCTTGCTCCCAAGTGTGGATTACAGCAGCAGCTGCG 22159

RESULT 31
ABA20511/C
ID ABA20511 standard; DNA; 29329 BP.

XX
AC ABA20511;
XX
XX 23-JAN-2002 (first entry)
DT
XX
DE Human nervous system related polynucleotide SEQ ID NO 12842.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;

KW antiparkinsonian; antisickling; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

OS Homo sapiens.
XX
XX WO200159063-A2.
XX
PD 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249247.
PR 17-NOV-2000; 2000US-0249248.
PR 17-NOV-2000; 2000US-0249249.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX

PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225271.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX Disclosure; SEQ ID NO 12842; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
XX (AB114678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 29329 BP; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 29329;
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccaccttgccctcccaaatgctggattacagggcatgagccactgcy 1238
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 22210 CTGCCACCTTGCCCTCCCAAACTGCTGGATTACAGGCATGAGCCACTGCG 22159

RESULT 32

AAK70791/C
 ID AAK70791 standard; DNA; 29329 BP.

XX AC AAK70791;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25603.
 XX KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytotstatic; gene therapy; vaccine; metastasis; ds.
 XX OS Homo sapiens.

XX PN WO2001571182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.

PR 12-SEP-2000; 2000US-0231960.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483426/52.
 DR
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 33324; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 29329 BP; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 29329;
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1187 ctgcccaccttgccctcccaaaagtgcgtggattacaggatgagccactgcg 1238
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 22210 CTGCCACCTTGGCCTCCCAAGTCTGGGATTACAGCATGAGCCACTGCG 22159
 RESULT 34
 AAC25949/C
 ID AAC25949 standard; cDNA; 149 BP.
 XX
 AC AAC25949;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 30024.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.

```

XX PF 21-FEB-2000; 2000EP-0200610.
XX DR
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 30024; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
XX Sequence 149 BP; 37 A; 43 C; 40 G; 28 T; 1 other;

Query Match 1.6%; Score 51; DB 21; Length 149;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttgacctcccaagtgctgggattacaggcatgagccactgcccagc 1244
Db 66 CTTGGGCTCCCAAGTGCTGGGATTACAGCATGAGCCACTGGGCCAGC 16

RESULT 35
AAI81653/C
ID AAI81653 standard; cDNA; 396 BP.
XX
XX AAI81653;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 1713.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO08027.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -

WPI; 2001-514838/56.
P-PSDB; AAO01722.
Isolated nucleic acids and polypeptides, useful for preventing
diagnosing and treating e.g. leukaemia, inflammation and immune
disorders -
Claim 1; SEQ ID NO 1713; 1399pp + Sequence Listing; English.
The invention relates to human polynucleotides (AAI79941-AAI93841) and
the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 396 BP; 124 A; 75 C; 87 G; 110 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccaccttgccctcccaagtgctgggattacaggcatgagccactgc 1237
Db 378 CTGCCACCTTGGCTCCCAAGTGCTGGGATTACAGGCATGAGCCACTGC 328

RESULT 36
AAI87958/C
ID AAI87958 standard; cDNA; 425 BP.
XX
XX AAI87958;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 8018.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO08027.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -

```

XX PS Claim 1; SEQ ID NO 8018; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO1910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 425 BP; 141 A; 83 C; 84 G; 116 T; 1 other;

Query Match 1.68; Score 51; DB 22; Length 425;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 ccttgccctcccaagctgggtacagcagcatgagccactgcgccagc 1244
|||||
Db 174 CTTGGCCTCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAGC 124

RESULT 37
AAK77921
ID AAK77921 standard; DNA; 1856 BP.
XX AC AAK77921;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32733.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -

XX Disclosure; SEQ ID NO 37733; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.

XX Sequence 1856 BP; 375 A; 416 C; 358 G; 707 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 1856;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1187 ctgccacccttgccctcccaagctgctggattacagcagcactgc 1237
|||||
Db 740 ctgccacccttgccctcccaagctgctggattacagcagcactgc 790

RESULT 38
AAI64545/c

ID AAI64545 standard; cDNA; 2046 BP.

XX AC AAI64545;

XX DT 23-NOV-2001 (first entry)

XX Human polypeptide-cytochrome b5-13 encoding cDNA.

XX Human; polypeptide-cytochrome b5-13; malignant tumour; haemopathy; HIV;
KW human immunodeficiency virus; infection; immunological disease;
KW inflammation; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1083..1451

XX /*tag= a

XX /product= "polypeptide-cytochrome b5-13"
XX /note= "claimed in claim 6"

XX CML301705-A.

XX PD 04-JUL-2001.

XX PF 27-DEC-1999; 99CN-0125352.

XX PR 27-DEC-1999; 99CN-0125352.

XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX WPI; 2001-550466/62.

XX P-PSDB; AAG78149.

XX New polypeptide-cytochrome b 5-13 and its encoding polynucleotide
PT useful for treating tumour, immunological disease, haemopathy, human
PT immunodeficiency virus infection and inflammation -

XX Claim 6; Page 23-24 (Disclosure); 30pp; Chinese.

XX The invention relates to the human polypeptide-cytochrome b5-13 and its
CC encoding polynucleotide. The polypeptide is used to treat various
CC diseases, such as malignant tumour, haemopathy, human immunodeficiency
CC virus infection, immunological diseases and various inflammation.

XX Sequence 2046 BP; 608 A; 472 C; 411 G; 555 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 2046;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 cttggcctcccaagctgctggattacagcagcactgcgccagc 1244
|||||
Db 1542 CCTGGCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCCAGC 1492

RESULT 39
AAK77916

ID AAK77916 standard; DNA; 7759 BP.
XX AC AAK77916;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32728.
XX DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 3728; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 7759 BP; 1935 A; 1488 C; 1467 G; 2869 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 7759;
 Best Local Similarity 100.0%; Pred. No. 4.4e-13;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacacttgccctcccaaaagtgtggtgattacaggatgagccactgc 1237
 |||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 5097 ctgccacacttgccctcccaaaagtgtggtgattacaggatgagccactgc 5147

RESULT 40
 ABA07913/C
 ID ABA07913 standard; DNA; 17904 BP.
 XX
 AC ABA07913;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human ovarian and breast cancer associated polynucleotide SEQ ID NO 708.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ds.
 XX
 OS Homo sapiens.
 OS
 XX WO200155325-A2.
 PN
 XX
 XX 02-AUG-2001.
 PD

XX 17-JAN-2001; 2001WO-US01345.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-488786/53.

PT New isolated ovarian and/or breast cancer related nucleic acids and
 PT polypeptides, useful for diagnosing, treating and/or preventing human
 PT diseases and disorders, particularly ovarian and/or breast cancer -
 XX Disclosure; SEQ ID NO 708; 577pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA07454-ABA08224) and proteins
 CC (ABB10743-ABB10980) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_sequences.
 XX
 SQ Sequence 17904 BP; 5485 A; 3834 C; 3913 G; 4672 T; 0 other;
 Query Match 1.6%; Score 51; DB 22; Length 17904;
 Best Local Similarity 100.0%; Pred. No. 4.4e-13;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1190 cccacettggcccccacgaagctggattacagccatgagccactgcgcc 1240
 Db 4315 CCCACCTGGCCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCC 4265
 RESULT 41
 AAL03730/G
 ID AAL03730 standard; DNA; 17904 BP.
 XX AC AAL03730;
 XX AC AAL03730;
 XX 21-NOV-2001 (first entry)
 XX Human reproductive system related antigen DNA SEQ ID NO: 6418.
 DE Human reproductive system related antigen; reproductive system disorder;
 KW Human; cancer; gene therapy; ds.
 KW Human; cancer; gene therapy; ds.
 XX Homo sapiens.
 XX WO200155320-A2.
 XX 02-AUG-2001.
 XX 17-JAN-2001; 2001WO-US01339.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214896.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225737.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen -
is used in preventing, treating or ameliorating a medical condition -

Disclosure; SEQ ID NO 6418; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a
number of human reproductive system related antigens. These can be used
in the prevention and treatment of reproductive system disorders,
including cancer. The present sequence is a genomic sequence encoding a
protein of the invention.

Sequence 17904 BP; 5485 A; 3834 C; 3913 G; 4672 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 17904;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 ccacattgctcccaaaagtctgggtattacaggcatgagccactgccc 1240
AA
Db 4315 CCACCTTGCCCTCCCAAAAGTCTGGATTACAGGCATGAGCCACTGCGCC 4265

RESULT 42
AAL36313/c
ID AAL36313 standard; DNA; 26591 BP.
XX AAL36313;
AC
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2678.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
XX WO200155367-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209457.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.
XX 14-SEP-2000; 2000US-0232401.
XX 14-SEP-2000; 2000US-0233063.
XX 14-SEP-2000; 2000US-0233064.
XX 14-SEP-2000; 2000US-0233065.
XX 21-SEP-2000; 2000US-0234223.
XX 21-SEP-2000; 2000US-0234274.
XX 25-SEP-2000; 2000US-0234997.
XX 25-SEP-2000; 2000US-0234998.
XX 26-SEP-2000; 2000US-0234984.
XX 27-SEP-2000; 2000US-0235834.
XX 27-SEP-2000; 2000US-0235836.
XX 29-SEP-2000; 2000US-0236327.
XX 29-SEP-2000; 2000US-0236367.
XX 29-SEP-2000; 2000US-0236368.
XX 29-SEP-2000; 2000US-0236369.
XX 29-SEP-2000; 2000US-0236370.
XX 02-OCT-2000; 2000US-0236802.
XX 02-OCT-2000; 2000US-0237037.
XX 02-OCT-2000; 2000US-0237038.
XX 02-OCT-2000; 2000US-0237039.
XX 02-OCT-2000; 2000US-0237040.
XX 13-OCT-2000; 2000US-0239935.
XX 13-OCT-2000; 2000US-0239937.
XX 20-OCT-2000; 2000US-0240960.
XX 20-OCT-2000; 2000US-0241221.
XX 20-OCT-2000; 2000US-0241785.
XX 20-OCT-2000; 2000US-0241786.
XX 20-OCT-2000; 2000US-0241787.
XX 20-OCT-2000; 2000US-0241808.
XX 20-OCT-2000; 2000US-0241809.
XX 20-OCT-2000; 2000US-0241826.
XX 01-NOV-2000; 2000US-0244617.
XX 08-NOV-2000; 2000US-0246474.
XX 08-NOV-2000; 2000US-0246475.
XX 08-NOV-2000; 2000US-0246476.
XX 08-NOV-2000; 2000US-0246477.
XX 08-NOV-2000; 2000US-0246478.
XX 08-NOV-2000; 2000US-0246523.
XX 08-NOV-2000; 2000US-0246524.
XX 08-NOV-2000; 2000US-0246525.
XX 08-NOV-2000; 2000US-0246526.
XX 08-NOV-2000; 2000US-0246527.
XX 08-NOV-2000; 2000US-0246528.
XX 08-NOV-2000; 2000US-0246532.
XX 08-NOV-2000; 2000US-0246609.
XX 08-NOV-2000; 2000US-0246610.
XX 08-NOV-2000; 2000US-0246611.
XX 08-NOV-2000; 2000US-0246613.
XX 17-NOV-2000; 2000US-0249207.
XX 17-NOV-2000; 2000US-0249208.
XX 17-NOV-2000; 2000US-0249209.
XX 17-NOV-2000; 2000US-0249210.
XX 17-NOV-2000; 2000US-0249211.
XX 17-NOV-2000; 2000US-0249212.
XX 17-NOV-2000; 2000US-0249213.
XX 17-NOV-2000; 2000US-0249214.
XX 17-NOV-2000; 2000US-0249215.
XX 17-NOV-2000; 2000US-0249216.
XX 17-NOV-2000; 2000US-0249217.
XX 17-NOV-2000; 2000US-0249218.
XX 17-NOV-2000; 2000US-0249219.

PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 XX WPI; 2001-451937/48.
 XX

XX Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -
 XX

PS Example 2; SEQ ID NO 2678; 781pp + Sequence Listing; English.
 XX

CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 26591 BP; 7489 A; 6354 C; 5953 G; 6795 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 26591;
 Best Local Similarity 100.0%; Pred. No. 4.4e-13;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacaccttgccctcccaagtgctggattacagcagcactgc 1237
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 11281 CTGCCACCTTGGCTCCCAAGTGTGGGATTACAGCAGCCTGC 11231

RESULT 43
 AAS34422
 ID AAS34422 standard; DNA; 32186 BP.
 XX
 AC AAS34422;
 XX
 DT 17-DEC-2001 (first entry)

XX
 DE Human DNA for a novel foetal antigen, SEQ ID No 1846.
 XX
 XX Human; foetal tissue antigen; ds; antiinflammatory; neuroprotective;
 KW immunomodulator; cardiovascular; cytostatic; nephrothropic;
 KW cardiovascular; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; breast neoplasm; cancer;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; epithelial cell proliferation; food additive.
 XX
 OS Homo sapiens.
 XX
 XX WO200155312-A2.
 PN
 XX
 XX 02-AUG-2001.
 PD
 XX
 XX 17-JAN-2001; 2001WO-US01321.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.


```

OS Homo sapiens.
PN WO200026382-A1.
XX 11-MAY-2000.
XX 28-OCT-1999; 99WO-JP05983.
XX 30-OCT-1998; 98JP-0310422.
XX (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
XX Toji S, Yano M, Tamai K;
XX WPI; 2000-365627/31.
XX Thiorodoxin reductase II (an X-linked inhibitor for apoptosis protein
XX (XIAP)-binding protein), with thiorodoxin reductase activity, useful
XX e.g. for treating apoptosis-related disorders, cancer and inflammation
XX PT
XX
XX Claim 15; Page 74-135; 139pp; Japanese.
XX
XX This sequence represents a human thiorodoxin reductase (TxR) encoding
XX genomic nucleotide sequence. The invention relates to two TxR proteins,
XX which have XIAP (X-linked inhibitor of apoptosis protein)-binding protein
XX activity. The human TxR gene is located on chromosome 22q11.2. The
XX invention includes antibodies which bind to the proteins, a vector
XX containing the TxR encoding nucleotide sequences, and methods for
XX producing transformants using the vector. Thiorodoxin reductase has
XX cytostatic and anti-inflammatory activity, and is used in the treatment
XX of diseases relating to apoptosis particularly due to cancer of viral
XX infection. TxR can also be used to treat inflammation and in the
XX screening of anti-cancer agents.
XX
XX Sequence 66566 BP; 13694 A; 17724 C; 18303 G; 16844 T; 1 other;

Query Match 1.6%; Score 51; DB 21; Length 66566;
Best Local Similarity 100.0%; Pred. No. 4.3e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttggcctcccaagctgctggattacagcagcagcactgcccagc 1244
|||||
Db 1791 ccttggcctcccaagctgctggattacagcagcagcactgcccagc 1841

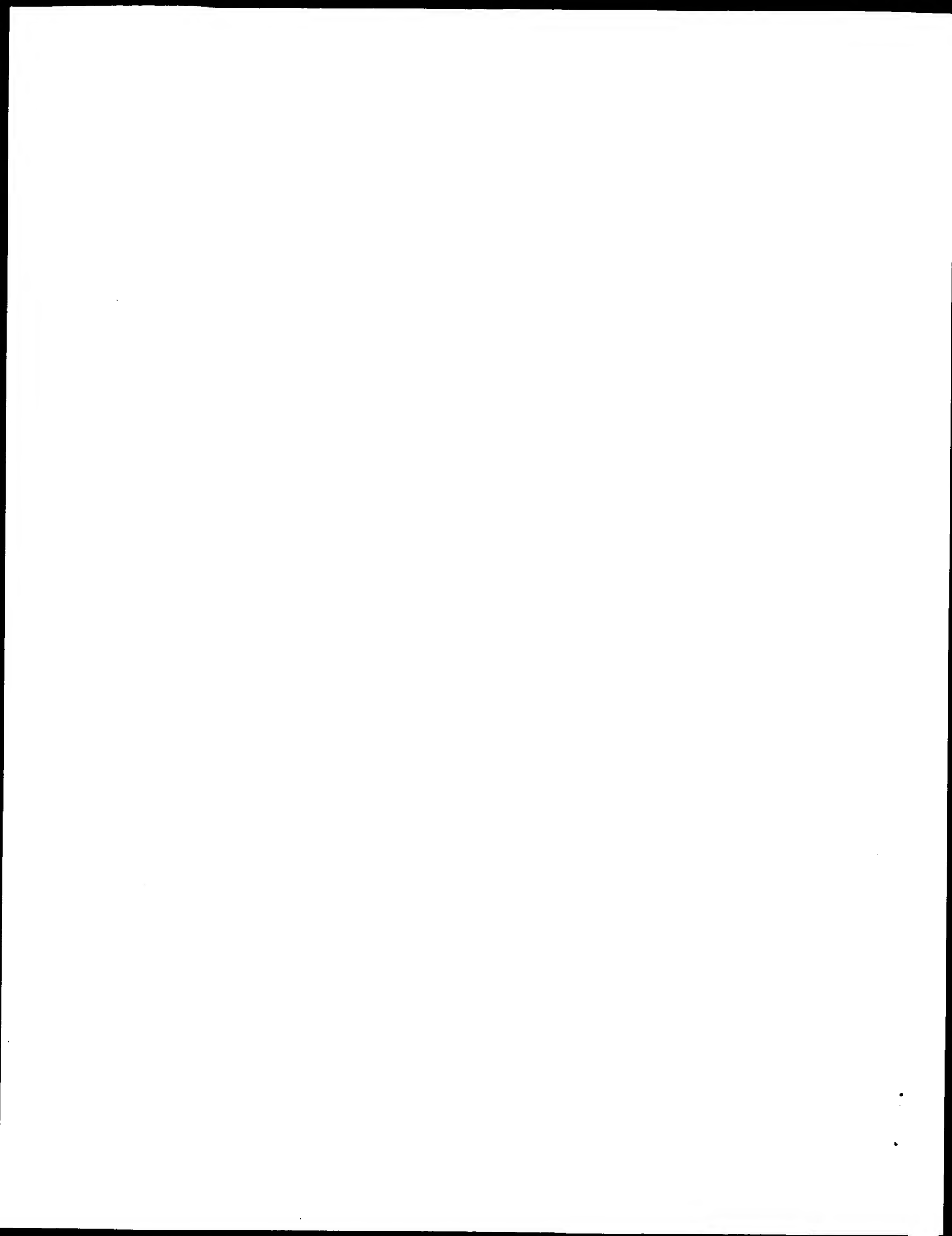
RESULT 45
AAC89560
ID AAC89560 standard; DNA; 122186 BP.
XX
XX AAC89560;
XX
XX 08-MAR-2001 (first entry)
XX
XX Human histone deacetylase HDAC-D coding sequence.
XX
XX Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
XX HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;
XX gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200071703-A2.
XX
XX 30-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-IB01252.
XX
XX 03-MAY-1999; 99US-0132287.
XX
XX (METH-) METHYLGENE INC.
XX
PI Macleod AR, Li Z, Besterman JM;
XX
XX WPI; 2001-016407/02.
XX
XX Antisense oligonucleotide that inhibits expression of a histone
XX deacetylase, useful for treating and/or alleviating the symptoms of
XX neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX
XX Disclosure; Page 89-125; 125pp; English.
XX
XX The present invention provides inhibitors of histone deacetylase enzymes
XX such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
XX inhibitors may be antisense strands or they may be compounds identified
XX by contacting the enzyme with the compound and measuring the resulting
XX enzyme activity. These inhibitors are useful for treating cancers and for
XX identifying which histone deacetylase is involved in a neoplasia.
XX
XX Sequence 122186 BP; 29016 A; 31077 C; 32425 G; 29668 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 122186;
Best Local Similarity 100.0%; Pred. No. 4.3e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttggcctcccaagctgctggattacagcagcagcactgcccagc 1244
|||||
Db 65150 ccttggcctcccaagctgctggattacagcagcagcactgcccagc 65200

Search completed: September 20, 2002, 06:06:55
Job time: 10299 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:09:41 ; Search time 111.56 Seconds
(without alignments)
7114.036 Million cell updates/sec

Title: US-09-846-456-1
Perfect score: 3231
Sequence: 1 acaggcatgtgagcagtg.....gccccacatccccaccatt 3231

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	55	1.7	246240	2 US-08-724-394A-20	Sequence 20, Appl
C 2	55	1.7	246240	2 US-08-724-394A-21	Sequence 21, Appl
C 3	55	1.7	246240	2 US-08-724-394A-22	Sequence 22, Appl
C 4	48	1.5	31571	1 US-08-323-443B-1	Sequence 1, Appl
C 5	48	1.5	53526	3 US-08-658-136-2	Sequence 2, Appl
C 6	48	1.5	53577	3 US-08-658-136-1	Sequence 1, Appl
C 7	47	1.5	13158	2 US-08-687-080-105	Sequence 105, App
C 8	45	1.4	153	2 US-08-849-701-2	Sequence 2, Appl
C 9	45	1.4	891	4 US-09-247-155-141	Sequence 141, App
C 10	45	1.4	1701	4 US-09-078-294-9	Sequence 9, Appl
C 11	45	1.4	3267	2 US-08-257-963B-12	Sequence 12, Appl
C 12	45	1.4	3267	4 US-08-367-841A-12	Sequence 12, Appl
C 13	45	1.4	3267	5 PCT-US95-07201-12	Sequence 12, Appl
C 14	45	1.4	3663	4 US-09-499-884-11	Sequence 11, Appl
C 15	45	1.4	5581	4 US-08-973-544-1	Sequence 1, Appl
C 16	45	1.4	8409	4 US-09-167-681-37	Sequence 37, Appl
C 17	45	1.4	22481	4 US-08-367-841A-43	Sequence 43, Appl
C 18	45	1.4	22481	5 PCT-US95-07201-43	Sequence 43, Appl
C 19	45	1.4	84495	4 US-09-797-906-3	Sequence 3, Appl
C 20	45	1.4	246240	2 US-08-724-394A-20	Sequence 20, Appl
C 21	45	1.4	246240	2 US-08-724-394A-21	Sequence 21, Appl
C 22	45	1.4	246240	2 US-08-724-394A-22	Sequence 22, Appl
C 23	44	1.4	176373	3 US-09-128-155-17	Sequence 17, Appl
C 24	43	1.3	2310	1 US-08-471-570-9	Sequence 9, Appl
C 25	43	1.3	2676	1 US-08-471-570-7	Sequence 7, Appl
C 26	43	1.3	14796	4 US-08-975-080-35	Sequence 35, Appl
C 27	43	1.3	14796	4 US-09-630-706-10	Sequence 10, Appl

28 43 1.3 14796 4 US-09-496-694B-3 Sequence 3, Appl
29 43 1.3 28720 4 US-09-341-587-7 Sequence 7, Appl
30 43 1.3 72604 4 US-09-268-992-7 Sequence 7, Appl
31 42 1.3 1613 2 US-08-812-204-1 Sequence 1, Appl
32 42 1.3 1656 1 US-08-324-465-2 Sequence 2, Appl
33 42 1.3 1656 2 US-08-465-981-2 Sequence 2, Appl
34 42 1.3 1656 5 PCT-US93-11915-2 Sequence 2, Appl
35 42 1.3 1725 1 US-08-324-465-5 Sequence 5, Appl
36 42 1.3 1725 2 US-08-465-981-5 Sequence 5, Appl
37 42 1.3 1725 5 PCT-US93-11915-5 Sequence 5, Appl
38 42 1.3 2086 2 US-08-655-640-5 Sequence 5, Appl
39 42 1.3 5590 4 US-09-050-159-129 Sequence 129, App
40 41 1.3 198 2 US-08-967-101-107 Sequence 107, App
41 1.3 198 2 US-08-593-541-107 Sequence 107, App
42 41 1.3 198 3 US-09-124-698-107 Sequence 107, App
43 41 1.3 198 4 US-09-127-480-107 Sequence 107, App
44 41 1.3 198 4 US-08-496-841C-107 Sequence 107, App
45 41 1.3 386 2 US-08-967-101-103 Sequence 103, App

ALIGNMENTS

RESULT 1
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/323,443B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, S. Peter
; REGISTRATION NUMBER: 25,351
; REFERENCE/DOCKET NUMBER: 0372/0A462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: PKD1 GENOMIC
US-08-323-443B-1

Query Match 1.5%; Score 48; DB 1; Length 31571;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccaagtgctggattacagcatgagccac 1234
|||||
DB 1117 CTGCCACCTTGGCTCCCAAGTCTGGATTACAGCATGAGCCAC 1164

RESULT 5
US-08-658-136-2
; Sequence 2, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: ONE MOUNTAIN ROAD
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
; REGISTRATION NUMBER: 31,845
; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-658-136-2

Query Match 1.5%; Score 48; DB 3; Length 53526;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccaagtgctggattacagcatgagccac 1234
|||||
DB 1712 CTGCCACCTTGGCTCCCAAGTCTGGATTACAGCATGAGCCAC 1759

RESULT 6
US-08-658-136-1
; Sequence 1, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: ONE MOUNTAIN ROAD
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
; REGISTRATION NUMBER: 31,845
; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53577 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

Query Match      1.5%; Score 48; DB 3; Length 53577;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacctggcctcccaagtgctgggattacagcagcatgagccac 1234
      |||||||
Db 1712 CTGCCACCTTGGCTCCCAAGTGCTGGGATTACAGGCGATGAGCCAC 1759

```

```

RESULT 7
US-08-687-080-105
; Sequence 105, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,136
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13158 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 5' END OF INTRON 21 OF RAD50 GENOMIC
; INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-105

```

```

Query Match      1.5%; Score 47; DB 2; Length 13158;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1198 ggctcccaagtgctgggattacagcagcatgagccacgccccac 1244
      |||||||
Db 336 GGCTCCCAAGTGCTGGGATTACAGGCGATGAGCCACTGCCCCAGC 382

```

```

RESULT 8

```

```

US-08-849-701-2/c
; Sequence 2, Application US/08849701
; Patent No. 5922544
; GENERAL INFORMATION:
; APPLICANT: Miyai, Kiyoshi
; APPLICANT: Naitoh, Tsutomu
; APPLICANT: Yonekawa, Toshihiro
; TITLE OF INVENTION: Method of Cell Detection
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/02734
; FILING DATE: 27-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: EIKEN1.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: Alu sequence BLUR1
US-08-849-701-2

```

```

Query Match      1.4%; Score 45; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 ccacacttgccctcccaagtgctgggattacagcagcatgagccac 1234
      |||||||
Db 85 CCACCTTGGCTCCCAAGTGCTGGGATTACAGGCGATGAGCCAC 41

```

```

RESULT 9
US-09-247-155-141/c
; Sequence 141, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET 021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116

```

; EARLIER FILING DATE: 1998-08-10
 ; EARLIER APPLICATION NUMBER: 60/099,273
 ; EARLIER FILING DATE: 1998-10-04
 ; NUMBER OF SEQ ID NOS: 182
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 141
 ; LENGTH: 891
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4...447
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 4...147
 ; OTHER INFORMATION: Von Heijne matrix
 ; OTHER INFORMATION: score 5.69999980926514
 ; OTHER INFORMATION: seq LLLFFGKLLVVG/VG
 ; FEATURE:
 ; NAME/KEY: polyA_signal
 ; LOCATION: 858..863
 ; FEATURE:
 ; NAME/KEY: polyA_site
 ; LOCATION: 880..891
 ; US-09-247-155-141

Query Match 1.4%; Score 45; DB 4; Length 891;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 aaccttgcctccgggtcaagcattctcctcgcctcagcctcc 1077
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 772 AACCTCTCCTCCGGGTTCAGCGATTCTCCTCCTCAGCCCTCC 728

RESULT 10
 US-09-294-9
 ; Sequence 9, Application US/09078294
 ; Patent No. 6265211
 ; GENERAL INFORMATION:
 ; APPLICANT: Choo, Kong-Hong Andy
 ; APPLICANT: Du Sart, Desiree
 ; APPLICANT: Cancilla, Michael R.
 ; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
 ; FILE REFERENCE: Davies Col
 ; CURRENT APPLICATION NUMBER: US/09/078,294
 ; CURRENT FILING DATE: 1998-05-13
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 1701
 ; TYPE: DNA
 ; ORGANISM: BAC-F2 contig 5
 ; US-09-078-294-9

Query Match 1.4%; Score 45; DB 4; Length 1701;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 aaccttgcctccgggtcaagcattctcctcgcctcagcctcc 1077
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1182 aaccttgcctccgggtcaagcattctcctcgcctcagcctcc 1226

RESULT 11
 US-08-257-963B-12
 ; Sequence 12, Application US/08257963B
 ; Patent No. 5840686
 ; GENERAL INFORMATION:
 ; APPLICANT: Chader, Gerald J.; Becerra, S.
 ; APPLICANT: Patricia; Schwartz, Joan P.;

; APPLICANT: Taniwaki, Takayuki
 ; TITLE OF INVENTION: PIGMENT EPITHELIUM
 ; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
 ; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
 ; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morgan & Finnegan
 ; STREET: 345 Park Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy Disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/257,963B
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/952,796
 ; FILING DATE: 24-SEPT-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DOROTHY R. AUTH
 ; REGISTRATION NUMBER: 36434
 ; REFERENCE/DOCKET NUMBER: 20264126US1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3267 Base Pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Unknown
 ; MOLECULE TYPE: Genomic DNA
 ; FEATURE:
 ; NAME/KEY: JTI09
 ; LOCATION:
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION: 3.3 kb PCR product
 ; OTHER INFORMATION: using primers, SEQ ID No. 5840686 15 and 16
 ; US-08-257-963B-12

Query Match 1.4%; Score 45; DB 2; Length 3267;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 ccacattggctcccaagtgctggattacagcattagccac 1234
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 613 CCCACCTTGCCCTCCCAAGTGTGGATTACAGCATGAGCCAC 657

RESULT 12
 US-08-367-841A-12
 ; Sequence 12, Application US/08367841A
 ; Patent No. 6319687
 ; GENERAL INFORMATION:
 ; APPLICANT: Chader, Gerald J.; Rodriguez,
 ; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
 ; APPLICANT: Tombran-Tink, Joyce
 ; TITLE OF INVENTION: PIGMENT EPITHELIUM
 ; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
 ; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morgan & Finnegan
 ; STREET: 345 Park Avenue
 ; CITY: New York

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3267 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: JT109
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 3.3 kb PCR product
; OTHER INFORMATION: using primers, SEQ ID No: 15 and 16
PCT-US95-07201-12

Query Match 1.4%; Score 45; DB 5; Length 3267;
Best Local Similarity 100.0%; Pred. No. 1.5e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 45; Conservative 0;

QY 1190 ccaccttgccctcccaagtcgtggattacagcgtagccac 1234
      |||||||||||||||||||||||||||||||||||||||
Db 613 CCCACCTGGCCCTCCCAAGTCTGGGATTACAGGCATGAGCCAC 657

RESULT 14
US-09-499-884-11/c
; Sequence 11, Application US/09499884
; Patent No. 6265172
; GENERAL INFORMATION:
; APPLICANT: St. Clair, Daret
; APPLICANT: Urano, Muneyasu
; APPLICANT: Kasarskis, Edward
; TITLE OF INVENTION: DIAGNOSTIC TEST AND THERAPY FOR MANGANESE SUPEROXIDE DISMUTASE
; TITLE OF INVENTION: ASSOCIATED DISEASES
; FILE REFERENCE: 50229-180
; CURRENT APPLICATION NUMBER: US/09/499,884
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 3663
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-499-884-11

Query Match 1.4%; Score 45; DB 4; Length 3663;
Best Local Similarity 100.0%; Pred. No. 1.5e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 45; Conservative 0;

QY 1033 aacctgtgctcccggttcaagcgtattcctgcctcagctccc 1077
      |||||||||||||||||||||||||||||||||||||||

```


Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aaactctgctcccggttcaagcgattctctgctcagctcc 1077
|||||
Db 4362 AACCTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCC 4318

RESULT 16

US-09-167-681-37
; Sequence 37, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboum, M.D., Richard M.
; APPLICANT: Raftogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 8397
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3730)...(3879)
; NAME/KEY: CDS
; LOCATION: (3987)...(4112)
; NAME/KEY: CDS
; LOCATION: (4198)...(4293)
; NAME/KEY: CDS
; LOCATION: (6088)...(6213)
; NAME/KEY: CDS
; LOCATION: (6309)...(6404)
; NAME/KEY: CDS
; LOCATION: (7214)...(7393)
; NAME/KEY: CDS
; LOCATION: (7516)...(7629)
US-09-167-681-37

Query Match 1.4%; Score 45; DB 4; Length 8409;

Best Local Similarity 100.0%; Pred. No. 1.4e-10;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 aaactctgctcccggttcaagcgattctctgctcagctcc 1077
|||||
Db 701 aaactctgctcccggttcaagcgattctctgctcagctcc 745

RESULT 17

US-08-367-841A-43
; Sequence 43, Application US/08367841A
; Patent No. 6319687
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
; APPLICANT: Tombran-Tink, Joyce
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:

Db 1970 AACCTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCC 1926

RESULT 15

US-08-973-544-1/C
; Sequence 1, Application US/08973544
; Patent No. 6338950
; GENERAL INFORMATION:
; APPLICANT: WEISS, Elisabeth
; TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,544
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT EP 96/02663
; FILING DATE: 20-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95109511.6
; FILING DATE: 20-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95112201.9
; FILING DATE: 03-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitts, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8341-7073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 48..162
; FEATURE:
; NAME/KEY: exon
; LOCATION: 544..652
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1044..1162
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1475..1567
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1775..1797
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2325..2709
US-08-973-544-1

Query Match 1.4%; Score 45; DB 4; Length 5581;

Best Local Similarity 100.0%; Pred. No. 1.5e-10;

```

; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367,841A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Pl-147
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: full length genomic
; OTHER INFORMATION: sequence for PEDF plus flanking sequences.
;
US-08-367-841A-43

```

```

Query Match          1.4%; Score 45; DB 4; Length 22481;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtgtgggtacagcagcatgagccac 1234
|||||
Db 17174 CCCACCTTGGCCTCCCAAAGTCTGGGATTACAGGCATGAGCCAC 17218

```

```

RESULT 18
PCT-US95-07201-43
; Sequence 43, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Pl-147
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: full length genomic
; OTHER INFORMATION: sequence for PEDF plus flanking sequences.
;
PCT-US95-07201-43

```

```

Query Match          1.4%; Score 45; DB 5; Length 22481;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtgtgggtacagcagcatgagccac 1234
|||||
Db 17174 CCCACCTTGGCCTCCCAAAGTCTGGGATTACAGGCATGAGCCAC 17218

```

```

RESULT 19
US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO01151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
;
US-09-797-906-3

```

```

Query Match          1.4%; Score 45; DB 4; Length 84495;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 caccttggcctcccaaaagtgtgggtacagcagcatgagccactg 1236
|||||
Db 26148 caccttggcctcccaaaagtgtgggtacagcagcatgagccactg 26192

```

RESULT 20

US-08-724-394A-20

; Sequence 20, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cdna

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H.CONTIG"

; US-08-724-394A-20

Query Match

Best Local Similarity 1.4%; Score 45; DB 2; Length 246240;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033

aacctctgctccgggttaagcgattctctgctcagctcc 1077

|||||

Db 242021 AACCTCTGCTCCGGGTTCAGCGATTCTCTGCTCAGCTCC 242065

RESULT 21

US-08-724-394A-21

; Sequence 21, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cdna

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H.CONTIG"

; US-08-724-394A-21

Query Match

Best Local Similarity 1.4%; Score 45; DB 2; Length 246240;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033

aacctctgctccgggttaagcgattctctgctcagctcc 1077

|||||

Db 242021 AACCTCTGCTCCGGGTTCAGCGATTCTCTGCTCAGCTCC 242065

RESULT 22

US-08-724-394A-22

; Sequence 22, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-22

Query Match 1.4%; Score 45; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1033 aacctgtcctccgggtccaagcgattctcgtccacgctcc 1077
|||||
Db 242021 AACCTGTGCTCCGGGTTCGAAGGATTCTCCTGCCTCACCTCC 242065

RESULT 23
US-09-128-155-17/c
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 1.4%; Score 44; DB 3; Length 176373;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1194 ccttgccctcccaagctgctggattacagcagcactgc 1237
|||||

Db 45158 CCTTGGCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGC 45115
RESULT 24
US-08-471-570-9/c
Sequence 9, Application US/08471570
Patent No. 5750371
GENERAL INFORMATION:
APPLICANT: IGARASHI, Koichi
APPLICANT: SENOO, Masaharu
APPLICANT: WATANABE, Tatsuya
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESS: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,570
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/149,664
FILING DATE:
APPLICATION NUMBER: US 07/743369
FILING DATE: 16-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: LINEK, Ernest V
REGISTRATION NUMBER: 29822
REFERENCE/DOCKET NUMBER: 40897
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2310 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1980
US-08-471-570-9

Query Match 1.3%; Score 43; DB 1; Length 2310;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1198 ggcctcccaagctgctggattacagcagcactgcgcc 1240
|||||
Db 2066 GGCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGC GCC 2024

RESULT 25
US-08-471-570-7/c
Sequence 7, Application US/08471570
Patent No. 5750371
GENERAL INFORMATION:
APPLICANT: IGARASHI, Koichi
APPLICANT: SENOO, Masaharu
APPLICANT: WATANABE, Tatsuya

```

RESULT      26
US-08-975-080-35
; Sequence 35, Application US/08975080
; Patent No. 6245523
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

RESULT 28
US-09-496-694B-3

; Sequence 3, Application US/09496694B
; Patent No. 6335194
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Eric E. Swayze
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
; FILE REFERENCE: ISPH-0439
; CURRENT APPLICATION NUMBER: US/09/496,694B
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 09/286,407
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 09/163,162
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 249
; SEQ ID NO 3
; LENGTH: 14796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2811)...(2921)
; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; NAME/KEY: CDS
; LOCATION: (5158)...(5275)
; NAME/KEY: CDS
; LOCATION: (11955)...(12044)
US-09-496-694B-3

Query Match 1.3%; Score 43; DB 4; Length 14796;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 cctctgctcccggttcaagcgattctctgctcagctccc 1077
|||||
Db 2320 cctctgctcccggttcaagcgattctctgctcagctccc 2362

RESULT 29

US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 28720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-7

Query Match 1.3%; Score 43; DB 4; Length 28720;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aacctgctcccggttcaagcgattctctgctcagctccc 1075
|||||
Db 19480 aacctgctcccggttcaagcgattctctgctcagctccc 19522

RESULT 30

US-09-268-992-7/c

; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match 1.3%; Score 43; DB 4; Length 72604;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1199 gctctccaaagtctgggtacagcagcatgagccactgccc 1241
|||||
Db 63856 GCCTCCCAAAGTCTGGATTACAGCATGAGCCACTGGCC 63814

RESULT 31

US-08-812-204-1
; Sequence 1, Application US/08812204
; Patent No. 5965790
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: SR-BI REGULATORY SEQUENCES AND
; FILE REFERENCE: THERAPEUTIC METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,204
; FILING DATE: 06-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-014.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:

	Query Match	1.3%;	Score 42;	DB 1: Length 1656;
	Best Local Similarity	100.0%;	Pred. No. 3.2e-09;	
	Matches 42;	Conservative 0;	Mismatches 0;	Indels
QY	1040	gctcccggggttcaagcagattctctgctcagcctcctgag	1081	
Db	332	GCCTCCCGGGTTCAGCGATTCTCTGCTCAGCTCCTGAG	373	

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/999,742
; FILING DATE: December 31, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US93-11915-2

Query Match 1.3%; Score 42; DB 5; Length 1656;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctccgggttcaagcagattctctgcctcagcctcctgag 1081
|||||
Db 332 GCCTCCGGGTTCAGCGATTCTCTGCTCAGCCTCCTGAG 373

RESULT 35
US-08-324-465-5
; Sequence 5, Application US/08324465
; Patent No. 5565334
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald
; APPLICANT: Abe, Miyako
; TITLE OF INVENTION: GENE TRANSCRIPTION AND
; TITLE OF INVENTION: IONIZING RADIATION: METHODS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,465
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/999,742
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065001

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1725
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-324-465-5

Query Match 1.3%; Score 42; DB 1; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctccgggttcaagcagattctctgcctcagcctcctgag 1081
|||||
Db 332 GCCTCCGGGTTCAGCGATTCTCTGCTCAGCCTCCTGAG 373

RESULT 36
US-08-465-981-5
; Sequence 5, Application US/08465981
; Patent No. 5874415
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald
; APPLICANT: Abe, Miyako
; TITLE OF INVENTION: ENHANCER SEQUENCE FOR MODULATING
; TITLE OF INVENTION: EXPRESSION IN EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,981
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/324,465
; FILING DATE: October 17, 1994
; APPLICATION NUMBER: 07/999,742
; FILING DATE: December 31, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1725
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-465-981-5

Query Match 1.3%; Score 42; DB 2; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1040 gctcccggttcaagcattctcctcagcctcctgag 1081
Db 332 GCTCCCGGTTCAAGGATTCTCTGCTCAGCCTCCTGAG 373

RESULT 37
PCT-US93-11915-5
; Sequence 5, Application PC/TUS9311915
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald
; TITLE OF INVENTION: ENHANCER SEQUENCE FOR MODULATING
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/999,742
; FILING DATE: December 31, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1725
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US93-11915-5

Query Match 1.3%; Score 42; DB 5; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.le-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1040 gctcccggttcaagcattctcctcagcctcctgag 1081
Db 332 GCTCCCGGTTCAAGGATTCTCTGCTCAGCCTCCTGAG 373

RESULT 38
US-08-655-640-5
; Sequence 5, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOFERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARRY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON

Query Match 1.3%; Score 42; DB 5; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.le-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1040 gctcccggttcaagcattctcctcagcctcctgag 1081
Db 332 GCTCCCGGTTCAAGGATTCTCTGCTCAGCCTCCTGAG 373

RESULT 39
US-09-050-159-129
; Sequence 129, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Leif T
; APPLICANT: Andersson, Maria K
; APPLICANT: Linstrom, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE THEREOF
; FILE REFERENCE: 1248/1D042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; EARLIER FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 5590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
; US-09-050-159-129

Query Match 1.3%; Score 42; DB 4; Length 5590;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 ctctgctcccggttcaagcattctcctgctcagcctcc 1077
Db 1036 ctctgctcccggttcaagcattctcctgctcagcctcc 1077

RESULT 40
US-08-655-640-5
; Sequence 5, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOFERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARRY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON

Db 1837 ctctgctccgggttaagcgattctctgctcagctcc 1878

RESULT 40
US-08-967-101-107
; Sequence 107, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-967-101-107

Query Match 1.3%; Score 41; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 ccttgccctccaaagtctggattacagggatgagccac 1234
|||||
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGGCATGAGCCAC 52

RESULT 41
US-08-592-541-107
; Sequence 107, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; CITY: Boston
; STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/08/592,541
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-592-541-107

Query Match 1.3%; Score 41; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 ccttgccctccaaagtctggattacagggatgagccac 1234
|||||
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGGCATGAGCCAC 52

RESULT 42
US-09-124-698-107
; Sequence 107, Application US/09124698
; Patent No. 6117978
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/124,698
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs

```

; Patent No. 6210919
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby, PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,841C
; FILING DATE: 28-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehlner, Ph.D.
; REGISTRATION NUMBER: 35,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 107:
;
US-08-496-841C-107

Query Match 1.3%; Score 41; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0;

Qy 1194 ccttggcctcccaagtctgggattacagcgatgagccac 1234
|||||
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGCGATGAGCCAC 52

RESULT 45
US-08-967-101-103
; Sequence 103, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-09-127-480-107

Query Match 1.3%; Score 41; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0;

Qy 1194 ccttggcctcccaagtctgggattacagcgatgagccac 1234
|||||
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGCGATGAGCCAC 52

RESULT 44
US-08-496-841C-107
; Sequence 107, Application US/09127480
; Patent No. 6194153
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-09-127-480-107

Query Match 1.3%; Score 41; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0;

Qy 1194 ccttggcctcccaagtctgggattacagcgatgagccac 1234
|||||
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGCGATGAGCCAC 52

RESULT 43
US-09-127-480-107
; Sequence 107, Application US/09127480
; Patent No. 6194153
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ROMMENS, JOHANNA M
; FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-09-127-480-107

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-967-101-103

```

```

Query Match      1.3%; Score 41; DB 2; Length 386;
Best Local Similarity 100.0%; Pred. No. 9.3e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1194 ccttgccctcccaagtgctgggattacagcgatgagccac 1234
      |||||||||||||||||||||||||||||||||||
Db 12 CCTTGGCCTCCCAAGTGCTGGGATTACAGCGATGAGCCAC 52

```

```

Search completed: September 20, 2002, 06:14:47
Job time: 11106 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 00:17:46 ; Search time 3900.56 Seconds
(without alignments)
11180.103 Million cell updates/sec

Title: US-09-846-456-1
Perfect score: 3231
Sequence: 1 acagggcatgtggcagggtg.....gccccacatccccaccactt 3231

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.*

- 1: em_estha.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	336	10.4	763	9	AU121731
2	217	6.7	736	9	AU135588
3	154	4.8	292	10	244377 HSC12B081 n
4	126	3.9	289	9	AW748338 RC6-BT025
5	111	3.4	467	9	AA527406 ng37c02.s
6	98	3.0	998	10	BG578861 602624760
7	58	1.8	617	9	BE178833 PM0-HT060
8	57	1.8	219	12	AQ38994 RPC1-11-3
9	57	1.8	424	9	AA652813 ns74c05.s
10	57	1.8	674	12	AG067644 Pan trogl
11	56	1.7	569	10	BI059823 IL3-UT011
12	56	1.7	600	10	BI015195 MR4-ET014
13	56	1.7	1054	10	BG745912 602724092
14	55	1.7	384	10	BG014649 IL5-GN023
15	55	1.7	387	10	BG007260 IL5-GN023
16	55	1.7	444	12	AQ056873 CIT-HSP-2
17	54	1.7	325	9	AA010265 z109e05.s

18	54	1.7	405	10	BF431825
19	54	1.7	468	9	AW104031
20	54	1.7	814	10	BG576148
21	53	1.6	453	12	AQ663862 HS_2151_B
22	53	1.6	658	9	AI610607
23	52	1.6	233	10	BI062371
24	52	1.6	249	10	BI710918
25	52	1.6	348	10	BI711369
26	52	1.6	391	9	AI264119
27	52	1.6	414	9	AA481408
28	52	1.6	416	9	AI679442
29	52	1.6	417	9	AI337065
30	52	1.6	417	9	AI679952
31	52	1.6	450	10	BI062368
32	52	1.6	455	9	AI089524
33	52	1.6	484	9	AI002969
34	52	1.6	501	12	AQ489571
35	52	1.6	530	9	AW971724
36	52	1.6	545	12	AQ623696
37	52	1.6	583	9	AI084593
38	52	1.6	717	12	AQ589333
39	52	1.6	753	9	AV718287
40	51	1.6	171	12	AQ198714
41	51	1.6	232	9	AA662976
42	51	1.6	360	10	H90008
43	51	1.6	369	9	AA486970
44	51	1.6	375	10	BF924753
45	51	1.6	379	12	AQ280600

ALIGNMENTS

RESULT 1

AU121731
LOCUS AU121731 MAMMAL Homo sapiens cDNA clone MAMMA1000851 5', mRNA
DEFINITION AU121731 sequence.
ACCESSION AU121731 763 bp mRNA linear EST 19-OCT-2000
VERSION AU121731.1 GI:10936966
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 763)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

Location/Qualifiers
1..763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1000851"
/clone_lib="MAMMA1"
/tissue_type="mammary gland"
/note="Vector: pME18SFL3"

BASE COUNT 137 a 205 c 260 g 158 t 3 others
ORIGIN

Query Match 10.48; Score 336; DB 9; Length 763;
 Best Local Similarity 100.0%; Pred. No. 1.9e-126;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2896 aattgcagcagagtgagtgaggccgagccacagagccagagccctctctccc 2955
 Db 1 AATTGCGAGCAGAGTGAGTGGCGCGGAGCCCGAGAGCCGAGCCCTTCTCC 60

QY 2956 gggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3015
 Db 61 GGGCTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 120

QY 3016 ttgctctgttt 3075
 Db 121 TTTGCTCTGTTTTCCTCCGGTCTGTTTTCCTCCGGTCTGTTTTCCTCCGG 180

QY 3076 ggttaggaagagcagcagcagcagcagcagcagcagcagcagcagcagc 3135
 Db 181 GGTAGGAGAAAGACACCAACAAAGTGGAACAGAGTAAAGGCTCTCCAGTGACT 240

QY 3136 tacttgagcgttatgttttttttttttttttttttttttttttttttttt 3195
 Db 241 TACTTGGCGGCTATGTGTTTTCGAGGCAAGAGGCTTCGGAAGTCTCGGTTTCG 300

QY 3196 gggacttgatccggagccacatccaccactt 3231
 Db 301 GGGACTTGTATCCGGAGCCCATCCACCACTT 336

RESULT 2
 LOCUS AU135588 736 bp mRNA linear EST 24-OCT-2000
 DEFINITION AU135588 PLACE1 Homo sapiens CDNA clone PLACE1002437 5', mRNA
 sequence.
 ACCESSION AU135588
 VERSION AU135588.1 GI:10996127
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
 Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
 Isogai,T.
 TITLE HRI human CDNA project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3951
 Fax: 81-438-52-3952
 Email: genomics@hri.co.jp
 HRI human CDNA project; 5' - & 3'-end one pass sequencing; Helix
 Research Institute; CDNA library construction; Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute

FEATURES
 source
 1..736
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE1002437"
 /clone_lib="PLACE1"
 /tissue_type="placenta"
 /note="Vector: pME18SFL3"
 BASE COUNT 163 a 199 c 199 g 170 t 5 others
 ORIGIN

Query Match 6.7%; Score 217; DB 9; Length 736;
 Best Local Similarity 100.0%; Pred. No. 6.9e-78;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2898 ttgcagcagagtgagtgaggccgagccacagagccagagccctctctccc 2957
 Db 5 TTGCGAGCAGAGTGAGTGGCGCGGAGCCCGAGAGCCGAGCCCTTCTCCCGG 64

QY 2958 gctgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3017
 Db 65 GCTGCGCAGGCGAGGCGGAGGCTCCGCGCACCAACAGAGCGGTTCTCAGGCGCTT 124

QY 3018 tgcctctgttt 3077
 Db 125 TGTCTCTGTTTTCCTCCGGTCTGTTTTCCTCCGGTCTGTTTTCCTCCGG 184

QY 3078 taggaagagagcagcagcagcagcagcagcagcagcagcagcagcagc 3114
 Db 185 TAGGAGAAAGAGAGCGCAACACAAAGTGTGAAACAG 221

RESULT 3
 LOCUS Z44377 292 bp mRNA linear EST 14-NOV-1994
 DEFINITION HSC12B081 normalized infant brain CDNA Homo sapiens CDNA clone
 c-1zb08, mRNA sequence.
 ACCESSION Z44377
 VERSION Z44377.1 GI:573506
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 292)
 AUTHORS Aufray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
 ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
 Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
 Sebastiani-Kabaktchis,C. and Tessier,A.
 IMAGE: molecular integration of the analysis of the human genome
 and its expression
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 95277534
 CONTACT: Genethon
 Genexpress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read.
 Genexpress_library_idt: C; Genexpress_sequence_idt: ylc-1zb08
 Seq primer: (-21)MT3 universal.

FEATURES
 source
 1..292
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-1zb08"
 /clone_lib="normalized infant brain CDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site1: HindIII;
 Site2: NotI; sex:Female; dev stage=3 months old;
 isolate=muscular atrophy patient; tissue_type=total brain
 ; total mRNA was oligo-(dr) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the
 lafmid BA vector. Clone library from B.Soaes, Psychiatry
 Dept. Columbia University, USA. Normalization_method:
 Bento Soares, P.N.A.S in press"
 BASE COUNT 50 a 87 c 96 g 56 t 3 others
 ORIGIN

Query Match 4.8%; Score 154; DB 10; Length 292;
 Best Local Similarity 99.5%; Pred. No. 5.7e-52;


```

Db 1 TTTGGGCTCTCTCTCAATTATGAAGAGCAGTAGTTCCTCGGTCC 60
|||||
QY 2212 tctgaggacctgggagctcaggctgggaatctccaaggcagtaggtcgc 2262
|||||
Db 61 TCTGAGGACCTGGGAGCTCAGGCTGGGATCTCCAGGAGCAGTAGGTCC 111
|||||

RESULT 6
BG678861 998 bp mRNA linear EST 01-MAY-2001
LOCUS 602624760F1 NCI_CGAP_Skn4 Homo sapiens cdna clone IMAGE:4749735 5',
DEFINITION mRNA sequence.
ACCESSION BG678861
VERSION BG678861.1 GI:13910258
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cdna Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0603 row: 9 column: 16
High quality sequence stop: 860.
Location/Qualifiers
1. 998
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749735"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 285 a 233 c 244 g 236 t
ORIGIN

Query Match 3.0%; Score 98; DB 10; Length 998;
Best Local Similarity 100.0%; Pred. No. 2.2e-29;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3017 ttgtctctgtttttcccccgttctgttttcccttctccggaagctgttcaagg 3076
|||||
Db 105 TTGCTCTCTGTTTTCCTCCGGTCTGTTTCTCCCTTCTCCGGAAGCTGTCTCAAGG 164
|||||

QY 3077 gtaggagaagagacgacaaacacaaagtgtgaaacag 3114
|||||
Db 165 GTAGGAGAAAGAGACGCAACACAAAGTGGAAACAG 202
|||||

RESULT 7
BE178833/c 617 bp mRNA linear EST 22-JUN-2000
LOCUS PMO-HT0608-170300-001-f02 HT0608 Homo sapiens cdna, mRNA sequence.
DEFINITION
ACCESSION BE178833
VERSION BE178833.1 GI:8657985
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 219)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
, J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI-11-343L18.TV

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 617)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2-PMO-HT0608-170
300-001-f02&t3=2000-03-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 612.
Location/Qualifiers
1. 617
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0608"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 194 a 128 c 150 g 145 t
ORIGIN

Query Match 1.8%; Score 58; DB 9; Length 617;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccaccttggtcccaagtgctggattacagcagcactgcgcagc 1244
|||||
Db 247 CTGCCACCTTGGCTCCGAAAGTCTGGGATTACAGCATGAGCCACTGCCCGCAGC 190
|||||

RESULT 8
AQ538994 219 bp DNA linear GSS 19-MAY-1999
LOCUS RPCI-11-343L18.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-343L18
DEFINITION , DNA sequence.
ACCESSION AQ538994
VERSION AQ538994.1 GI:4869633
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 219)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
, J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI-11-343L18.TV

```


BASE COUNT 152 a 162 c 179 g 179 t 2 others
ORIGIN

Query Match 1.8%; Score 57; DB 12; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1025 ctcaactgtaacctgctcccggttcaagcgattctcctgctcagctcctcgag 1081
|||||
Db 219 CTCACCTGTAACTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCTCTGAG 163

RESULT 11
BI059823
LOCUS
DEFINITION IL3-UT0115-080101-378-H06 UT0115 Homo sapiens cDNA, mRNA sequence. EST 15-JUN-2001
ACCESSION BI059823
VERSION BI059823.1 GI:14467350
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 569)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ci=IL3&t2=IL3-UT0115-080101-378-H06&t3=2001-01-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 568.
Location/Qualifiers
1..569
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0115"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

BASE COUNT 175 a 114 c 119 g 161 t
ORIGIN

Query Match 1.7%; Score 56; DB 10; Length 569;
Best Local Similarity 100.0%; Pred. No. 4e-12;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 ccacattgctcccaagtctggttgattacagcatgagccactgcccagct 1245
|||||
Db 120 CCCACCTTGCCCTCCCAAGTCTGGGATTACAGCATGAGCCGCGCCAGCT 175

RESULT 12
BI015195/c
LOCUS
DEFINITION MR4-ET0140-190201-006-e12 ET0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI015195
VERSION BI015195.1 GI:14419266
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 600)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ci=MR4&t2=MR4-ET0140-190201-006-e12&t3=2001-02-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 72.
Location/Qualifiers
1..600
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0140"
/dev_stage="Adult"
/note="Organ: lung_tumor; Vector: puc18; Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

BASE COUNT 208 a 102 c 130 g 160 t
ORIGIN

Query Match 1.7%; Score 56; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 3.9e-12;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacattgctcccaagtctggttgattacagcatgagccactgcccagc 1244
|||||
Db 451 GCCCACCCTTGCCCTCCCAAGTCTGGGATTACAGCATGAGCCGCGCCAGC 396

RESULT 13
BG745912
LOCUS
DEFINITION BG745912 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850375 3', mRNA sequence.
ACCESSION BG745912
VERSION BG745912.1 GI:14056565
KEYWORDS EST.

SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1054)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: L1CML1691 row: h column: 24
High quality sequence stop: 841.
Location/Qualifiers
1..1054
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4850375"
/clone_lib="NIH_MGC_113"
/lab_host="DH10B (phage-resistant)"
/note="Organ: spleen; Vector: pORF7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 265 a 270 c 231 g 288 t
ORIGIN

Query Match 1.7%; Score 56; DB 10; Length 1054;
Best Local Similarity 100.0%; Pred. No. 2.9e-12;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 gccaccttgcctcccaaaagtctggattacagcagcatgacgcactgcgcacgc 1244
|||||
Db 188 GCCCACCTTGGCCCTCCCAAAAGTCTGGGATTACAGGCATGCGCCACG 243

RESULT 14
BG014649
LOCUS 384 bp mRNA linear EST 24-JAN-2001
DEFINITION IL5-GN0239-211200-342-g07 GN0239 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG014649
VERSION BG014649.1 GI:12456063
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0239-211200-342-g07&t3=2000-11-27&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 371.

FEATURES
source
1..384
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0239"
/dev_stage="Adult"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 86 a 110 c 87 g 100 t
ORIGIN

Query Match 1.7%; Score 55; DB 10; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 ccacacttgcctcccaaaagtctggattacagcagcatgacgcactgcgcacgc 1244
|||||
Db 115 CCACCTTGGCCCTCCCAAAAGTCTGGGATTACAGGCATGCGCCACG 169

RESULT 15
BG007260
LOCUS 387 bp mRNA linear EST 24-JAN-2001
DEFINITION IL5-GN0239-271100-281-e06 GN0239 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG007260
VERSION BG007260.1 GI:12451267
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 387)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0239-271100-281-e06&t3=2000-11-27&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 371.


```

Db 194 CCCACCTTGCCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCCCG 247

RESULT 18
BF431825
LOCUS
DEFINITION
    BF431825 405 bp mRNA linear EST 19-JAN-2001
    nab50h08.x1 Soares NSF_F8_9W_OT_PA_P.S1 Homo sapiens cDNA clone
    IMAGE:3269606 3' similar to contains Alu repetitive element
    ;contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION
BF431825
VERSION
BF431825.1 GI:11443939
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 405)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 385.
FEATURES
    source
        1..405
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3269606"
            /clone_lib="Soares.NSF_F8_9W_OT_PA_P.S1"
            /lab_host="DH10B"
            /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
            a modified polylinker; Site_1: Not I; Site_2: Eco RI;
            Equal amounts of plasmid DNA from five normalized
            libraries were mixed, and ss circles were made in vitro.
            Following HAP purification, this DNA was used as tracer in
            a subtractive hybridization reaction. The driver was
            PCR-amplified cDNAs from pools of 5,000 clones made from
            the same 5 libraries. The pools consisted of the following
            libraries and cloneIDs: Soares NBHSF pool 1:
            309384-310919, 323208-325895 Soares Nb2HP pool 1:
            145032-147335, 147720-148103, 148872-149255, 15002 -
            150407, 151176-152327 Soares Nb2Hr8-9W pool 1:
            758280-760583, 772104-774407 Soares NbHrPA pool 1:
            304776-306311, 320136-322823, 326280-326663 Soares NBHOT
            pool 1: 723720-726407, 739080-740999 Subtraction by Bento
            Soares and M. Fatima Bonaldo."
BASE COUNT 83 a 113 c 84 g 125 t
ORIGIN

Query Match 1.7%; Score 54; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.1e-11;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccaccttgccctcccaagtgctgggattacagcagcatgagccactgccccg 1243
|||||
Db 233 CCCACCTTGCCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCCCG 286

RESULT 19
AW104031
LOCUS
DEFINITION
    AW104031 468 bp mRNA linear EST 20-OCT-1999
    x64a05.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2598512 3'
    similar to contains Alu repetitive element; mRNA sequence.
ACCESSION
AW104031
VERSION
AW104031.1 GI:6074766
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 468)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 424.
FEATURES
    source
        1..468
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:2598512"
            /clone_lib="NCI_CGAP_Ov23"
            /tissue_type="tumor, 5 pooled (see description)"
            /lab_host="DH10B"
            /note="Organ: ovary; Vector: PCMV-SPORT6; Site_1: Sall;
            Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.35 kb. Tumor types include: mixed
            Mullerian tumor, papillary serous, clear cell, spindle
            cell. All are primary tumors, metastasis positive. Life
            Technologies catalog #: 11534-013"
BASE COUNT 101 a 116 c 98 g 153 t
ORIGIN

Query Match 1.7%; Score 54; DB 9; Length 468;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccaccttgccctcccaagtgctgggattacagcagcatgagccactgccccg 1243
|||||
Db 250 CCCACCTTGCCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCCCG 303

RESULT 20
BG576148/c
LOCUS
DEFINITION
    BG576148 814 bp mRNA linear EST 10-APR-2001
    602597119F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4705901 5',
    mRNA sequence.
ACCESSION
BG576148
VERSION
BG576148.1 GI:13583801
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 814)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTp
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10573 row: e column: 06
High quality sequence stop: 779.

```

FEATURES
source

Location/Qualifiers
1. .814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4705901"
/clone_lib="NIH_MGC.87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
240 a 208 c 197 g 169 t

BASE COUNT
ORIGIN

Query Match 1.7%; Score 54; DB 10; Length 814;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1190 cccacctggctcccaaaagtctgggtacagcagcatgagccactgcgccag 1243
|||||
Db 172 CCCACCTTGGCTCCCAAAGTCTGGGATTACAGGCGATGAGCCACTGCCGCCAG 119

RESULT 21
A0663862

LOCUS
DEFINITION HS.2151.B1.G10.T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2151 Col-19 Row=N, DNA sequence.

ACCESSION A0663862
VERSION A0663862.1
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 453)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 2151 row: N column: 19
Seq primer: T7
Class: BAC ends
High quality sequence stop: 453.

Location/Qualifiers
1. .453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-2151 Col-19 Row=N"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
E-Coli DH10B"
108 a 114 c 82 g 149 t

FEATURES
source

Query Match 1.6%; Score 53; DB 12; Length 453;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1189 gcccacctggctcccaaaagtctgggtacagcagcatgagccactgcgcc 1241
|||||
Db 243 GCCCACCTTGGCTCCCAAAGTCTGGGATTACAGGCGATGAGCCACTGCCGCC 295

BASE COUNT
ORIGIN

Best Local Similarity 100.0%; Pred. No. 7.5e-11;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gcccacctggctcccaaaagtctgggtacagcagcatgagccactgcgcc 1241
|||||
Db 142 GCCCACCTTGGCTCCCAAAGTCTGGGATTACAGGCGATGAGCCACTGCCGCC 194

RESULT 22
A1610607

LOCUS
DEFINITION tp13g09.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2188288 3', similar to TR:Q99634 Q99634 RIG-G. [1]; contains Alu repetitive element; mRNA sequence.

ACCESSION A1610607
VERSION A1610607.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 658)
NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 2270 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 372
POLYA-No.

JOURNAL
COMMENT

Location/Qualifiers
1. .658
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2188288"
/clone_lib="NCI_CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with
signed ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. primer: Oligo dT.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"
130 a 196 c 144 g 187 t 1 others

FEATURES
source

Query Match 1.6%; Score 53; DB 9; Length 658;
Best Local Similarity 100.0%; Pred. No. 6.2e-11;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1189 gcccacctggctcccaaaagtctgggtacagcagcatgagccactgcgcc 1241
|||||
Db 243 GCCCACCTTGGCTCCCAAAGTCTGGGATTACAGGCGATGAGCCACTGCCGCC 295

BASE COUNT
ORIGIN

Query Match 1.6%; Score 53; DB 9; Length 658;
Best Local Similarity 100.0%; Pred. No. 6.2e-11;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1189 gcccacctggctcccaaaagtctgggtacagcagcatgagccactgcgcc 1241
|||||
Db 243 GCCCACCTTGGCTCCCAAAGTCTGGGATTACAGGCGATGAGCCACTGCCGCC 295

RESULT 23
B1062371

LOCUS
DEFINITION IL3-UT0117-160301-502-E10_1 UT0117 Homo sapiens cDNA, mRNA
sequence.
ACCESSION B1062371
VERSION B1062371.1
GI:14469898

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 233)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0117-
160301-502-E10.1&t3=2001-03-16&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 233.

FEATURES
Location/Qualifiers
1..233
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0117"
/dev_stage="Adult"
/note="Organ: uterus; tumor: Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
53 a 65 c 53 g 62 t
BASE COUNT
ORIGIN

Query Match 1.6%; Score 52; DB 10; Length 233;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 gccacacttgctcccaagtgctggattacagcagcactgccc 1240
|||||
Db 137 GCCCACCTTGGCTCCCAAGTGGTGGATTACAGGATGAGCCACTGCGCC 188
|||||

RESULT 24
BI710918/c 249 bp mRNA linear EST 19-SEP-2001
LOCUS id94g04.y1 Human insulinoma Homo sapiens cDNA 5', similar to
DEFINITION SW:ALU7_HUMAN P39194 ALU SUBFAMILY SQ SEQUENCE CONTAMINATION
WARNING ENTRY. [1] ;, mRNA sequence.

ACCESSION BI710918.1 GI:15686613
VERSION BI710918.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 249)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Schmitt, A., Theising, B., Rutter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Schmitt, A., Theising, B., Rutter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 190.

FEATURES
Location/Qualifiers
1..249
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human insulinoma"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1:
XhoI; Site: 2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

BASE COUNT 69 a 59 c 57 g 64 t
ORIGIN

Query Match 1.6%; Score 52; DB 10; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 gccacacttgctcccaagtgctggattacagcagcactgccc 1240
|||||
Db 120 GCCCACCTTGGCTCCCAAGTGGTGGATTACAGGATGAGCCACTGCGCC 69
|||||

RESULT 25
BI711369 348 bp mRNA linear EST 19-SEP-2001
LOCUS id94g04.x1 Human insulinoma Homo sapiens cDNA 3', similar to
DEFINITION SW:ALU7_HUMAN P39194 ALU SUBFAMILY SQ SEQUENCE CONTAMINATION
WARNING ENTRY. [1] ;, mRNA sequence.

ACCESSION BI711369
VERSION BI711369
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 348)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Rutter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs: id94g04.y1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 316.

FEATURES

source

1. .348

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Human insulinoma"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1:

XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system

(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue following the Washington

University protocol

(http://genome.wustl.edu/est/lambda_protocol.shtml).

Please contact Hiroshi Inoue, MD/PhD for further

information on this library (Metabolism Division, Permutt

Laboratory, Washington University School of Medicine, Box

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this

is a Washington University Pancreas EST project library."

80 a 89 c 79 g 100 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 1.6%; Score 52; DB 10; Length 348;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 gccacctggctcccaaaagtgtggattacagcgatgacccactgcgc 1240

|||||

Db 246 GCCCACCCTGGCTCCCAAGTGTGGGATTACAGCGATGACCCACTCGCC 297

RESULT 26

A1264119/c

LOCUS

DEFINITION q108h11.x1 Soares_NHMPu.s1 Homo sapiens cDNA clone IMAGE:1855941

3', similar to contains Alu repetitive element,, mRNA sequence.

ACCESSION A1264119

VERSION A1264119.1

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 391)

NCT-CGAP

http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1507 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .391

/organism="Homo sapiens"

/db_xref="taxon:9606"

FEATURES

source

1. .391

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1855941"

/clone_lib="Soares_NHMPu.s1"

/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"

/lab_host="DH10B"

/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 115 a 94 c 100 g 82 t

ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 391;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1193 acctggctcccaaaagtgtggattacagcgatgacccactgcgcgcgc 1244

|||||

Db 347 ACCTTGGCTCCCAAGTGTGGGATTACAGCGATGACCCACTCGCCGCC 296

RESULT 27

AA481408

LOCUS zV02q09.s1 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:746368 3'

DEFINITION similar to contains Alu repetitive element,, mRNA sequence.

ACCESSION AA481408

VERSION AA481408.1

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 414)

NCI-CGAP

http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41ml3 fwd. ET from Amersham

High quality sequence stop: 400.

Location/Qualifiers

1. .414

/organism="Homo sapiens"

/db_xref="GDB:5945900"

/db_xref="taxon:9606"

/clone="IMAGE:746368"

/clone_lib="NCI-CGAP GCBI"

/tissue_type="germinal center B cell"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-GTGTACCAATCTGAAGTGGAGCGCGCCATCTTTTTTTTTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 106 a 112 c 86 g 110 t

Query Match 1.6%; Score 52; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacctggctcccaagtgtggattacagcagcatgagccactgcgcc 1241
|||||
Db 180 CCCACCTGGCTCCCAAGTGTGGATTACAGGATGAGCCACGCGGCC 231
|||||

RESULT 28
AI679442
LOCUS
DEFINITION
tu63h07.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2255773 3', similar to contains Alu repetitive element; contains element TAR1 repetitive element ;, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
AI679442
AI679442.1 GI:4889624
EST.
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 416)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: Greg Lennon, Ph.D.
CDNA Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1074 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 400.

FEATURES
source
1. .416
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2255773"
/clone_lib="NCI_CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with signet ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 83 a 106 c 102 g 125 t

Query Match 1.6%; Score 52; DB 9; Length 416;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacctggctcccaagtgtggattacagcagcatgagccactgcgcc 1241
|||||
Db 210 CCCACCTGGCTCCCAAGTGTGGATTACAGGATGAGCCACGCGGCC 261
|||||

RESULT 29
AI337065

LOCUS
DEFINITION
qx82g11.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2009060 3', similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
AI337065
AI337065.1 GI:4073992
EST.
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 417)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 528 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 400.

FEATURES
source
1. .417
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2009060"
/clone_lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 101 a 97 c 94 g 125 t

ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 417;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacctggctcccaagtgtggattacagcagcatgagccactgcgcc 1240
|||||
Db 323 GCCCACCTGGCTCCCAAGTGTGGATTACAGGATGAGCCACGCGCC 374
|||||

RESULT 30
AI679952

LOCUS
DEFINITION
tu67h07.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2256157 3', similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION
VERSION
AI679952
AI679952.1 GI:4890134

```

KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         1 (bases 1 to 417)
JOURNAL       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.lnli.gov/bbrp/image/image.html
              Seq primer: -40UP from Gibco
              High quality sequence stop: 409.
FEATURES      Location/Qualifiers
              1..417
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="NCI-CGAP_Gas4"
              /tissue_type="poorly differentiated adenocarcinoma with
              signet ring cell features"
              /lab_host="DH10B"
              /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
              Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
              Average insert size 1.69 kb. Life Technologies catalog #:
              11549-011"
BASE COUNT    83 a 106 c 102 g 126 t
ORIGIN

Query Match      1.6%; Score 52; DB 9; Length 417;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacttgccctccaaagtctggattacagcagcatgagccactgcgcc 1241
|||||
Db 210 CCCACTTGGCTCCCAAAGTCTGGGATTACAGGCATGAGCCACTGCGCC 261

RESULT 31
BI062368      450 bp      mRNA      linear      EST 15-JUN-2001
LOCUS         IL3-UT0117-160301-502-G10 UT0117 Homo sapiens cDNA, mRNA sequence.
DEFINITION    BI062368
ACCESSION     BI062368
VERSION       BI062368.1 GI:14469895
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         1 (bases 1 to 450)
JOURNAL       Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
              Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
              Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
              Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
              M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
              Simpson, A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
              20202663
              Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research

FEATURES      source
              1..455
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="IMAGE:1676040"
              /dev_stage="8-9 weeks"
              /lab_host="DH10B"
              /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
              polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
              was prepared from mRNA obtained from pooled 8-9 week

```

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0117-160301-502-G10&t3=2001-03-16&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 450.

```

FEATURES      Location/Qualifiers
              1..450
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="UT0117"
              /dev_stage="Adult"
              /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the pUC 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
              low stringency conditions."
BASE COUNT    111 a 122 c 97 g 120 t
ORIGIN

```

```

Query Match      1.6%; Score 52; DB 10; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacttgccctccaaagtctggattacagcagcatgagccactgcgcc 1240
|||||
Db 157 GCCACCTTGGCTCCCAAAGTCTGGGATTACAGGCATGAGCCACTGCGCC 208

```

```

RESULT 32
AI089524      455 bp      mRNA      linear      EST 01-OCT-1998
LOCUS         OZ22a01.x1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION    IMAGE:1676040 3' similar to contains Alu repetitive element;; mRNA
              sequence.
ACCESSION     AI089524
VERSION       AI089524.1 GI:3428583
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         1 (bases 1 to 455)
JOURNAL       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Insert Length: 761 Std Error: 0.00
              Seq primer: -40ml3 fwd. ET from Amersham
              High quality sequence stop: 444.
              Location/Qualifiers
              1..455
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="IMAGE:1676040"
              /dev_stage="8-9 weeks"
              /lab_host="DH10B"

```

```

FEATURES      source
              1..455
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="IMAGE:1676040"
              /dev_stage="8-9 weeks"
              /lab_host="DH10B"
              /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
              polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
              was prepared from mRNA obtained from pooled 8-9 week

```

(total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCAAGTGGAGCGCCGCTTAATTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 136 a 105 c 111 g 103 t
ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1193 acctggcctcccaagtgtcggtacagcagcagcagcagcagcagc 1244
|||||
Db 359 ACCTTGGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGC 308

RESULT 33
LOCUS AI002969
DEFINITION an23g03.s1 Gessler Wilms tumor Homo sapiens cDNA clone
IMAGE:1699540 3' similar to contains Alu repetitive element
; contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION AI002969
VERSION AI002969.1 GI:3203383
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 484)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 448.

FEATURES
source
1..484
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1699540"
/clone_lib="Gessler Wilms tumor"
/sex="pooled (6)"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; RNA
was prepared from a pool of 6 anonymous Wilms' tumor RNAs.
RNA was prepared by acid-phenol, followed by one round of
oligo dT selection. cDNA library preparation was with
the BR/Life tech. Superscript Plasmid system. An
oligo-dT NotI primer for first strand synthesis generated
gcggccgcct(n) at the 3' end of the clones. A 5' SalI
adaptor was used with sequence 5'-gtcaccacgcgtcg-3'.
Resulting cDNAs were size selected (average size 2 kb),
NotI digested, and ligated into NotI/SalI-cut pSPORT1.
Library was constructed by Dr. Manfred Gessler."

BASE COUNT 126 a 131 c 103 g 124 t
ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 ccaccttgctcccaagtgtcggtacagcagcagcagcagcagcagc 1241
|||||
Db 227 CCCACTTGGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCC 278

RESULT 34
LOCUS AQ489571
DEFINITION RPCI-11-230P5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-230P5,
DNA sequence.

ACCESSION AQ489571
VERSION AQ489571.1 GI:4675445
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 501)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building

JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI-11-230P5.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbest@igf.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES
source
1..501
/organism="Homo sapiens"
/db_xref="GBB:7588300"
/db_xref="taxon:9606"
/clone="RPCI-11-230P5"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

BASE COUNT 106 a 123 c 107 g 165 t
ORIGIN

Query Match 1.6%; Score 52; DB 12; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1193 acctggcctcccaagtgtcggtacagcagcagcagcagcagcagc 1244
|||||
Db 395 ACCTTGGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGC 446

RESULT 35
LOCUS AW971724
DEFINITION EST383813 MAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION AW971724
VERSION AW971724.1 GI:8161570

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 530)
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt,
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 294
Seq primer: Forward.
Location/Qualifiers
1..530
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGL"
/note="Vector: pBluescriptSKm"
BASE COUNT 135 a 139 c 111 g 145 t
ORIGIN
Query Match 1.6%; Score 52; DB 9; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1190 cccacgtggctcccaagtgctggattacagcagcatgagccactgcgcctc 1241
|||||
Db 227 CCCAGCTGGCTCCCAAAGTCTGGGATTACAGGCATGAGCCACTGGGCC 278
|||||

RESULT 36
A0623696
LOCUS HS_5319_A2_C10_T7A RPCI-11 Human Male BAC Library GSS 16-JUN-1999
DEFINITION genomic clone Plate=895 Col=20 Row=E, DNA sequence.
ACCESSION A0623696
VERSION A0623696.1 GI:5086011
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3687
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 895 row: E column: 20

Seq primer: T7
Class: BAC ends
High quality sequence stop: 545.
Location/Qualifiers
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=895 Col=20 Row=E"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"
BASE COUNT 117 a 139 c 139 g 144 t 6 others
ORIGIN
Query Match 1.6%; Score 52; DB 12; Length 545;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1194 ccttggtcccaagtgctggattacagcagcatgagccactgcgcctc 1245
|||||
Db 349 CCTTGGCTCCCAAAGTCTGGGATTACAGGCATGAGCCACTGGCCCACT 400
|||||

RESULT 37
A1084593/c
LOCUS ox63f01.s1 Soares_NhMHPu_S1 Homo sapiens cDNA clone IMAGE:1661017
DEFINITION 3' similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION A1084593
VERSION A1084593.1 GI:3423016
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 470 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 411.
Location/Qualifiers
1..583
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1661017"
/clone_lib="Soares_NhMHPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pTT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 165 a 138 c 139 g 140 t 1 others

```



```

source      1. 171
/organism="Homo sapiens"
/db_xref="GDB:7522510"
/db_xref="taxon:9606"
/clone="RPCI-11-59723"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 43 a 32 c 44 g 52 t
ORIGIN

Query Match 1.6%; Score 51; DB 12; Length 171;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccaccttgctcccaagtctgggtacagcagcagcactgcgcc 1240
|||||
DB 100 CCCACCTTGCTCCCAAGTCTGGGTACAGCAGCAGCAGCAGC 50

RESULT 41
AA662976 232 bp mRNA linear EST 12-NOV-1997
LOCUS ac52b01.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
DEFINITION IMAGE:866185 3' similar to contains Alu repetitive element/, mRNA
sequence.
ACCESSION AA662976
VERSION AA662976.1 GI:2616967
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 232)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,X., Wylie,T., Waterston,R. and Wilson,R.
WASHU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. EF from Amersham
High quality sequence stop: 216.
FEATURES
Location/Qualifiers
1. 232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:866185"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
vector; -5' adaptor sequence: 5' GAATCGGCACGAG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 48 a 51 c 55 g 78 t
ORIGIN

Query Match 1.6%; Score 51; DB 9; Length 232;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1031 gtaacctgctcccggttcaagcgattctctgctgaagcctctctgag 1081
|||||
DB 115 GTAACCTCTGCTCCCGGTTCAAGCGATTCTCTGCTGCTCAGCCTCTCTGAG 165

RESULT 42
H90008 Homo sapiens
LOCUS H90008/c 360 bp mRNA linear EST 28-NOV-1995
DEFINITION YU83d03.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:240389 3' similar to contains Alu repetitive element; contains
MER26 repetitive element ;, mRNA sequence.
ACCESSION H90008
VERSION H90008.1 GI:1080438
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 360)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,I., Soares,M., Tan,F., Trevisakis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1095
High quality sequence stops: 320
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1095 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 320.
FEATURES
Location/Qualifiers
1. 360
/organism="Homo sapiens"
/db_xref="GDB:3789342"
/db_xref="taxon:9606"
/clone="IMAGE:240389"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAGAGATTATTAAGATCTTTTTTTTTTTTTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."
BASE COUNT 108 a 70 c 81 g 98 t 3 others
ORIGIN

Query Match 1.6%; Score 51; DB 10; Length 360;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttggcctcccaagctgctgggtacagcagcagcactgcgccagc 1244
|||||
DB 174 CCTTGGCCTCCCAAGTCTGGGATTACAGGATGAGCCACTGCGCCAGC 124

RESULT 43

```


Email: mdadamsetigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES

SOURCE

Location/Qualifiers

1..379
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2516H22"
 /clone_lib="CITBI-E1"
 /sex="male"
 /cell_type="sperm"
 /note="Vector: pBelBAC11; Site_1: EcoRI; Site_2: EcoRI;
 Caltech Human BAC Library D"

BASE COUNT
 ORIGIN

85 a 98 c 111 g 85 t

Query Match

Best Local Similarity 1.6%; Score 51; DB 12; Length 379;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 cctggcctcccaagctggtgattacagcagccactgcccagc 1244
 |||||
 DB 174 CCTTGGCTCCCAAGTCTGGGATTACAGGCGTGGCCGCGCCAGC 224

Search completed: September 20, 2002, 04:07:18
 Job time: 13772 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:25:43 ; Search time 5250.46 Seconds
(without alignments) 1422.880 Million cell updates/sec

US-09-846-456-2
 Title: 357
 Perfect score: 357
 Sequence: 1 tggaggtctcagctgagagg.....gagggaaggagctgtgttg 357

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

post-processing: Listing first 45 summaries

Database : GenEmbl : *

1: gb.ba:*
2: gb.hug:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.ro:*
10: gb.roi:*
11: gb.st:*
12: gb.sy:*
13: gb.un:*
14: gb.yi:*
15: en.ba:*
16: en.fun:*
17: en.hum:*
18: en.in:*
19: en.mu:*
20: en.om:*
21: en.or:*
22: en.ov:*
23: en.pat:*
24: en.ph:*
25: en.pl:*
26: en.ro:*
27: en.sts:*
28: en.un:*
29: en.vi:*
30: en.hug.h.hug.i
31: en.hug.i
32: en.hug.o
33: en.hug.o

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	ID	Description
NO					

357	100.0	357	6	AX351030	Sequence
255	71.4	149034	9	AF275948	Homo sapi
159	44.5	159	6	AX351033	Sequence
159	44.5	480	9	HSAX52277	Homo sapi
159	44.5	129608	9	AL353685	Human DNA
159	44.5	175064	2	AC012230	Homo sapi
159	44.5	183999	9	AX092589	Sequence
159	44.5	201144	9	AF287262	Homo sapi
116	32.5	90698	2	AC021345	Homo sapi
99	27.7	99	6	AX351035	Sequence
99	27.7	99	6	AX351036	Sequence
81	22.7	200	9	AF238622S2	Homo sapi
77	21.6	10442	6	AX060713	Sequence
77	21.6	10442	6	AX060892	Sequence
77	21.6	10442	9	AF285167	Homo sapi
77	21.6	10474	6	AX060719	Sequence
77	21.6	10474	6	AX060721	Sequence
77	21.6	10474	6	AX060898	Sequence
77	21.6	10474	6	AX060900	Sequence
77	21.6	10474	6	AB037924	Homo sapi
60	16.8	298	6	AX037924	Homo sapi
60	16.8	446	6	X127764	Sequence
60	16.8	446	6	AX139751	Sequence
60	16.8	697	9	AF258627	Homo sapi
60	16.8	6786	9	AB055982	Homo sapi
60	16.8	7260	6	AX253452	Sequence
60	16.8	7860	6	AX092594	Sequence
60	16.8	7862	6	AX135712	Sequence
60	16.8	9741	6	AX127830	Sequence
60	16.8	9741	6	AX139817	Sequence
60	16.8	9741	6	AX351036	Sequence
60	16.8	9854	6	AX127831	Sequence
60	16.8	9854	6	AX139818	Sequence
51	14.3	1556	9	AK024328	Homo sapi
35	9.8	37	6	AX092843	Sequence
35	9.8	38	6	AX092841	Sequence
27	7.6	207659	2	AC091466	Mus muscu
27	7.6	278572	10	AF287263	Mus muscu
24	6.7	12425	6	AX239607	Sequence
24	6.7	18427	9	AC006326	Homo sapi
24	6.7	90582	9	AL583882	Human DNA
24	6.7	123339	2	AF179296	Homo sapi
24	6.7	127270	9	AC036144	Homo sapi
24	6.7	132741	9	AL158043	Human DNA
24	6.7	134334	10	AL163512	Mouse DNA
24	6.7	137072	2	AF1353893	Homo sapi

ALIGNMENTS

```

RESULT      1
LOCUS      AX351030
DEFINITION Sequence 2 from Patent WO0183746.
ACCESSION  AX351030
VERSION    AX351030.1  GI:18616386
KEYWORDS   .
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (sites)
AUTHORS    Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,
            Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE      Regulatory nucleic acid sequences of the abcl gene
JOURNAL    Patent: WO 0183746-A 2 08-NOV-2001;
            Aventis Pharma S.A. (FR)
FEATURES   Location/Qualifiers
            1..357
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 64 a 78 c 116 g 99 t
ORIGIN

```

Query Match 100.0%; Score 357; DB 6; Length 357;
 Best Local Similarity 100.0%; Pred. No. 3.3e-197;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggaggtctcagctgagagggctggattagcagctctcattggtgtatgcttgcagca 60
 |||||
 Db 1 TGGAGGTCTCAGCTGAGAGGGCTGGATTAGCAGCTCTCATTTGGTGTATGGCTTTCAGCA 60
 |||||
 QY 61 ataactgatgctgtttccctcctcctctcttattcttctcatttgcagccacggcgt 120
 |||||
 Db 61 ATAACTGATGGCTGTTCCTCTCTCTTTATCTTTTCACTTAATGACCAACGCGGCT 120
 |||||
 QY 121 cctgtgtcagctctggcgtgccttcagggtcccgagccacacagcagctggcgtgct 180
 |||||
 Db 121 CCCTGTGTGAGCTGTGGCGGTGCTTCCAGGGTCCCGAGCCACAGCTGGCGGTGCT 180
 |||||
 QY 181 ggtgagggaacatgcatgttggcctcagctgaggtgctgtggaagaacctcact 240
 |||||
 Db 181 GGCTGAGGAACATGCGATGTGGCTCAGCTGAGTGTCTGCTGTGGAAGAACCTCACT 240
 |||||
 QY 241 ttcaagaagaacacagctagctggttttccagcagcgggggttctctctatcttt 300
 |||||
 Db 241 TTCAGAGAGACACACAGTAAGCTTGGGTTTTTCAGCAGCGGGGGTCTCTCAATTTT 300
 |||||
 QY 301 tcttgtggtttgagttgggagtgaggagggagggaggaagagcgtgtgtg 357
 |||||
 Db 301 TCTTTGTGTTTGTAGTTGGGATTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 357
 |||||

RESULT 2

AF275948 149034 bp DNA linear PRI 17-JUL-2000
 LOCUS Homo sapiens ABCAL (ABCAL) gene, complete cds.
 DEFINITION AF275948
 ACCESSION AF275948
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 149034)
 Santamarina-Fojo, S., Peterson, K., Knapper, C., Olu, Y., Freeman, L.,
 Cheng, J. F., Osorio, J., Remaley, A., Yang, X. P., Haudenschild, C.,
 Prades, C., Chimini, G., Blackmon, E., Francois, T., Duverger, N.,
 Rubin, E. M., Rosier, M., Deneffe, P., Fredrickson, D. S. and Brewer, H. B.
 Jr.

Complete genomic sequence of the human ABCAL gene: analysis of the
 human and mouse ATP-binding cassette A promoter

Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)

20345099

2 (bases 1 to 149034)

Santamarina-Fojo, S., Peterson, K. M., Knapper, C. L., Freeman, L. A.,

Remaley, A. T., Yang, X. P., Haudenschild, C. C., Blackmon, E. E.,

Francois, T. L. and Brewer, H. B. Jr.

Submitted (08-JUN-2000) Molecular Disease Branch, National

Institutes of Health, National Heart, Lung and Blood Institute,

Bethesda, MD 20892, USA

Location/Qualifiers

1. 149034

/organism="Homo sapiens"

/db_xref="taxon:9606"

join(1454, 1674, 25831, 25989, 40385, 40478, 45012, 45153,
 46423, 46541, 67729, 67850, 70831, 71007, 83960, 84052,
 89010, 89250, 91962, 92101, 92433, 92549, 96758, 96955,
 97702, 97907, 98428, 98604, 100391, 100613, 102360, 102581,
 103642, 103846, 104951, 105064, 106862, 107033,
 108023, 108154, 109460, 109602, 109806, 109943,
 110646, 110866, 112124, 112196, 113183, 113229,
 115053, 115101, 115297, 115410, 116806, 116954,
 118604, 118728, 119964, 120062, 123094, 123283,

124804, 124898, 126208, 126240, 127363, 127468,
 128943, 129017, 129539, 129708, 130939, 131116,
 133114, 133229, 133341, 133485, 134526, 134649,
 135736, 135865, 136130, 136250, 137038, 137100,
 138008, 138114, 140469, 140610, 140982, 141116,
 142060, 142163, 142646, 142738, 143397, 143640,
 144581, 148034)
 /gene="ABCAL"
 /product="ABCAL"
 join(1454, 1674, 25831, 25923)
 /gene="ABCAL"
 1454, 148034
 /gene="ABCAL"
 2406, 2678
 /rpt_family="Alu"
 4240, 4509
 /rpt_family="Alu"
 4998, 5270
 /rpt_family="Alu"
 5583, 5839
 /rpt_family="Alu"
 6615, 6877
 /rpt_family="Alu"
 8800, 9046
 /rpt_family="Alu"
 9998, 10279
 /rpt_family="Alu"
 11859, 12133
 /rpt_family="Alu"
 12810, 12902
 /rpt_family="Alu"
 15220, 15399
 /rpt_family="Alu"
 15508, 15787
 /rpt_family="Alu"
 18600, 18860
 /rpt_family="Alu"
 20950, 21206
 /rpt_family="Alu"
 25038, 25314
 /rpt_family="Alu"
 join(25924, 25989, 40385, 40478, 45012, 45153, 46423, 46541,
 67729, 67850, 70831, 71007, 83960, 84052, 89010, 89250,
 91962, 92101, 92433, 92549, 96758, 96955, 97702, 97907,
 98428, 98604, 100391, 100613, 102360, 102581, 103642, 103846,
 104951, 105064, 106862, 107033, 108023, 108154,
 109460, 109602, 109806, 109943, 110646, 110866,
 112124, 112196, 113183, 113385, 115053, 115101,
 115297, 115410, 116806, 116954, 118604, 118728,
 119964, 120062, 123094, 123283, 124804, 124898,
 126208, 126240, 127363, 127468, 128943, 129017,
 129539, 129708, 130939, 131116, 133114, 133229,
 13341, 133485, 134526, 134649, 135736, 135865,
 136130, 136250, 137038, 137100, 138008, 138114,
 140469, 140610, 140982, 141116, 142060, 142163,
 142646, 142738, 143397, 143640, 144581, 144721)
 /gene="ABCAL"
 /codon_start=1
 /product="ABCAL"
 /protein_id="AAF86276.1"
 /db_xref="GI:247086"

SYNOPSIS="MACWPQLLLKLNLTFRRTQCOLLEVAWPLFIFLILISRL
 SYPPQEHECFHPKAMPAGTLPWQGIICNANNPCFRYPTEAPGVGNKSVL
 RLSPDARLLYSQKDTSMKMRKVLRTLOQIKKSSNKLQDFLDVNETFSGFLVH
 ARSLPDKVDMRADVTLHKVFLQGLHLTSLCNGSKSEEMQLGQDEVELGLP
 REKLAAERVLRSNMILKPLTLTLNSTPSPSELAEATKLHLSLQGLAEFLSMR
 SWSMRQVEMFLTNWSSSTQIYQAVSRVCGHPGGGLKIKSLWYEDNNYKALF
 GNGTEADAETFDNSTTPPCNDLKNLESPLRIWKLKPLLVKILYTPPTAT
 QDMAEVNKTFOELAFHFDLEGMEELSPTWTWEMSCMDLVRMLDSRDNDHWE
 QQLDLWTAQDIAFLAKHPEDVQSSSVYTWREAFNETNOAIRTSREMECVNL
 KLEPIATEVWMLINKSMELLDERKFWAGIVFTGTPGSLPHVVKYKIRMDINVERT
 NKIKDGYMDPGPRADPFEDRMRYVNGFAYLDQVVEQAIIRVLTCTEKTGTVMQOMPY
 PCYVDDIFLRWMSRSMPLFTWLAIWYSVAVIKIGIYVEKEARLRETMRGLONSLW

QY 100 ttaatacagcagcagggcgctcgtctgctcagctctgcccgcgtcctccagggctccc 159
 Db 1 TTAATACAGCAGCAGCGGGCGTCCCTGCTCAGCTCTGCGCCGCTGCCTCCAGGGCTCCC 60
 QY 160 gagccacagctggcgctgctgaggaagaaatggaatgttgccctcagctgaggttg 219
 Db 61 GAGCCACAGCTGGGCGTCTGGCTGAGGGAACATGGCATGTGGCCCTCAGCTGAGGTG 120
 QY 220 ctgctgtggaagaacacctcacttccagaagaagacaaca 258
 Db 121 CTGCTGTGAAGAACCTCACTTTCAGAGAAGACACAACA 159
 RESULT 4
 HSA252277
 LOCUS Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, exon 2. linear PRI 08-JAN-2001
 DEFINITION
 ACCESSION AJ252277
 VERSION AJ252277.1 GI:12140344
 KEYWORDS ABC-1 gene; ATP-binding cassette transporter-1.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 Porsch-Oezcuernomez, M., Langmann, T. and Schmitz, G.
 Cloning and Characterization of the human ATP-binding Cassette Transporter-1 (ABC-1) Promoter
 Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 480)
 AUTHORS Porsch-Oezcuernomez, M.K.
 TITLE Direct Submission
 JOURNAL Submitted (07-JAN-2000) Porsch-Oezcuernomez M.K., Institute for Clinical Chemistry, University of Regensburg, Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY
 FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q31"
 /cell_type="leukocyte"
 189..346
 /gene="ABC-1"
 /number=2
 189..280
 /gene="ABC-1"
 /function="cholesterol efflux regulatory protein"
 189..346
 /gene="ABC-1"
 281..>346
 /gene="ABC-1"
 /function="cholesterol efflux regulatory protein"
 /codon_start=1
 /product="ATP-binding cassette transporter-1"
 /protein_id="CAC21428.1"
 /db_xref="GI:12140345"
 /translation="MACWPQLRLWKNLTERRQT"
 BASE COUNT 89 a 102 c 155 g 134 t
 ORIGIN
 Query Match 44.5%; Score 159; DB 9; Length 480;
 Best Local Similarity 100.0%; Pred. No. 2,1e-81;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcacttccagaagaagacaaca 258
 Db 287 TGTGGCCCTCAGCTGAGGTGCTGCTGGAAGAAGCTCACTTTCAGAGAAGACACAACA 346
 QY 259 gtaagcttggttttcagcagcgagggttctctctattttcttgggtttgagtt 318
 Db 259 gtaagcttggttttcagcagcgagggttctctctattttcttgggtttgagtt 318

Db 347 GTAAGCTGGCTTTTTCAGCAGCGGGGTTCTCTCATTTTCTTTTGTGCTTTGAGTT 406
 QY 319 ggggattggaaggaggaggaggaggaggaggagctgtgttg 357
 Db 407 GGGGATTGGAGGAGGAGGAGGAGGAGGAGGAGGAGCTGTGTG 445
 RESULT 5
 AL353685/c
 LOCUS Human DNA sequence from clone RP11-31J20 on chromosome 9, complete sequence. linear PRI 01-JUN-2001
 DEFINITION
 ACCESSION AL353685
 VERSION AL353685.23 GI:14329534
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 Tracey, A.
 Direct Submission
 Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jun 8, 2001 this sequence version replaced gi:14272260.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at:
 http://www.sanger.ac.uk/projects/Celegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-31J20 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACE3.6
 IMPORTANT: This sequence is not the entire insert of clone RP11-31J20. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
 The true right end of clone RP11-31J20 is at 129608 in this sequence. The true right end of clone RP11-413C10 is at 2000 in this sequence.
 Location/Qualifiers
 1..129608
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-31J20"
 /clone_lib="RPCI-11.1"
 1358..1653
 /note="AluSc repeat: matches 1..288 of consensus"
 1828..1877
 /note="L2 repeat: matches 2649..2698 of consensus"
 2496..2714
 /note="MIR repeat: matches 12..250 of consensus"
 2777..2896
 /note="MIR repeat: matches 6..128 of consensus"

repeat_region 3237. .3415
/note="LIME repeat: matches 5696. .5821 of consensus"
repeat_region 6522. .6818
/note="AluSq repeat: matches 1. .295 of consensus"
repeat_region 7282. .7415
/note="LIMB8 repeat: matches 6040. .6173 of consensus"
repeat_region 8145. .8434
/note="AluSc repeat: matches 1. .298 of consensus"
repeat_region 12145. .12713
/note="L2 repeat: matches 1363. .1940 of consensus"
repeat_region 13890. .13969
/note="L2 repeat: matches 2611. .2701 of consensus"
repeat_region 15380. .15411
/note="16 copies 2 mer ac 87% conserved"
repeat_region 16105. .16144
/note="10 copies 4 mer caca 100% conserved"
repeat_region 16868. .17049
/note="MIR repeat: matches 64. .246 of consensus"
repeat_region 17941. .18229
/note="AluSq repeat: matches 1. .287 of consensus"
repeat_region 18259. .18553
/note="AluSq repeat: matches 1. .293 of consensus"
repeat_region 20310. .20616
/note="AluSq repeat: matches 1. .308 of consensus"
repeat_region 20957. .21107
/note="MIR repeat: matches 49. .212 of consensus"
repeat_region 21783. .22078
/note="AluSc repeat: matches 7. .302 of consensus"
repeat_region 22320. .22439
/note="MIR repeat: matches 10. .146 of consensus"
repeat_region 22533. .22839
/note="AluSc repeat: matches 1. .307 of consensus"
repeat_region 23427. .23945
/note="L2 repeat: matches 2137. .2750 of consensus"
repeat_region 24245. .24544
/note="AluSq repeat: matches 2. .302 of consensus"
repeat_region 24556. .24587
/note="8 copies 4 mer acac 96% conserved"
repeat_region 26504. .26561
/note="29 copies 2 mer ta 69% conserved"
repeat_region 26849. .26892
/note="11 copies 4 mer tata 81% conserved"
repeat_region 27998. .28103
/note="MIR repeat: matches 28. .145 of consensus"
repeat_region 28515. .28626
/note="MIR repeat: matches 17. .129 of consensus"
repeat_region 28986. .29213
/note="MIR repeat: matches 2. .245 of consensus"
repeat_region 30121. .30422
/note="AluSq repeat: matches 1. .302 of consensus"
repeat_region 31424. .31734
/note="AluSp repeat: matches 1. .308 of consensus"
repeat_region 31987. .32116
/note="MER5A repeat: matches 60. .188 of consensus"
repeat_region 32450. .32536
/note="MER5A repeat: matches 14. .106 of consensus"
repeat_region 33293. .33380
/note="22 copies 4 mer atgg 79% conserved"
repeat_region 34435. .34588
/note="MER5A repeat: matches 13. .189 of consensus"
repeat_region 34729. .34873
/note="MIR repeat: matches 49. .198 of consensus"
repeat_region 35802. .35951
/note="MIR repeat: matches 77. .248 of consensus"
repeat_region 37183. .37260
/note="2 copies 39 mer 92% conserved"
repeat_region 37673. .37980
/note="AluSc repeat: matches 1. .308 of consensus"
repeat_region 39674. .40243
/note="LIMD2 repeat: matches 5774. .6331 of consensus"
repeat_region 40256. .40534
/note="L2 repeat: matches 2256. .2533 of consensus"
repeat_region 41476. .41615

repeat_region 42010. .42194
/note="MIR repeat: matches 30. .185 of consensus"
repeat_region 44809. .45239
/note="LTR16A repeat: matches 6. .450 of consensus"
repeat_region 45359. .45486
/note="MIR repeat: matches 1. .139 of consensus"
repeat_region 46654. .46693
/note="10 copies 4 mer acac 97% conserved"
repeat_region 47118. .47429
/note="AluSc repeat: matches 1. .312 of consensus"
repeat_region 47915. .48083
/note="AluSq repeat: matches 145. .313 of consensus"
repeat_region 49097. .49181
/note="L2 repeat: matches 2625. .2710 of consensus"
repeat_region 49578. .49758
/note="MIR repeat: matches 9. .188 of consensus"
repeat_region 49847. .49873
/note="MIR repeat: matches 155. .182 of consensus"
repeat_region 49874. .50188
/note="AluJb repeat: matches 3. .311 of consensus"
repeat_region 50189. .50347
/note="MIR repeat: matches 5. .155 of consensus"
repeat_region 51677. .51977
/note="AluSc repeat: matches 1. .303 of consensus"
repeat_region 52025. .52123
/note="L2 repeat: matches 1655. .1753 of consensus"
repeat_region 54337. .54544
/note="MIR repeat: matches 25. .246 of consensus"
repeat_region 54754. .55032
/note="AluSp repeat: matches 1. .302 of consensus"
repeat_region 55042. .55343
/note="AluY repeat: matches 5. .298 of consensus"
repeat_region 56027. .56232
/note="MIR repeat: matches 21. .259 of consensus"
repeat_region 56233. .56286
/note="18 copies 3 mer tgt 72% conserved"
repeat_region 56342. .56673
/note="L1MA9 repeat: matches 5893. .6307 of consensus"
repeat_region 56750. .56889
/note="L2 repeat: matches 2565. .2710 of consensus"
repeat_region 57406. .57567
/note="MER5A repeat: matches 4. .187 of consensus"
repeat_region 57864. .58174
/note="AluSq repeat: matches 1. .311 of consensus"
repeat_region 59443. .60004
/note="L2 repeat: matches 2088. .2723 of consensus"
repeat_region 60201. .60355
/note="MIR repeat: matches 20. .183 of consensus"
repeat_region 63991. .64159
/note="MIR repeat: matches 78. .243 of consensus"
repeat_region 64590. .64653
/note="32 copies 2 mer ta 70% conserved"
repeat_region 64593. .64656
/note="16 copies 4 mer atat 68% conserved"
repeat_region 64981. .65042
/note="MER5A repeat: matches 15. .72 of consensus"
repeat_region 65043. .65341
/note="AluSc repeat: matches 1. .299 of consensus"
repeat_region 65342. .65441
/note="MER5A repeat: matches 72. .186 of consensus"
repeat_region 66825. .67131
/note="AluSc repeat: matches 1. .306 of consensus"

Query Match 44.5%; Score 159; DB 9; Length 129608;
Best Local Similarity 100.0%; Pred. No. 2.7e-81;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 ttttgccagctgaggttctgtgtggaagacacctcactttcagaagaacaaca 258
|||||
Db 120383 TGTGTGGCCTCAGCTGAGGTGCTGCTGTGGAGAACCTCCTCTTTCAGAGAGAACAA 120324
|||||
Qy 259 gtaagcttggtttttcagcagcgggggtctctcatctttttttgtgtttgagtt 318

```

|||||
Db 120323 GTAAGCTGGGTTTTCAGCAGCGGGGCTCTCTCATTTTCTTGGTTTGAGTT 120264
QY 319 ggggattgagagggagggagggaggaagctgtgttg 357
|||||
Db 120263 GGGGATGGAGGAGGAGGAGGAGGAAGGAAGCTGTGTG 120225
|||||

RESULT 6
AC012230/c
LOCUS AC012230.3 GI:7637254
DEFINITION Homo sapiens clone RP11-IM10, WORKING DRAFT SEQUENCE, 39 unordered
          pieces.
ACCESSION AC012230
VERSION AC012230.3
KEYWORDS HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175064)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-IM10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175064)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
          Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckghalter,B.,
          Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
          Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
          Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
          Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
          Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
          Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
          McEwan,P., McGurk,A., McKernan,K., McDonald,P., Marquis,N.,
          Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
          Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
          Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
          Direct Submission
          Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          On Apr 22, 2000 this sequence version replaced gi:6454033.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          ----- Genome Center
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WIBR
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence_submissions@genome.wi.mit.edu
          ----- Project Information
          Center project name: L2510
          Center clone name: 1.M.10
          ----- Summary Statistics
          Sequencing vector: M13; M7815; 100% of reads
          Chemistry: Dye-terminator Big Dye; 100% of reads
          Assembly program: Phrap; version 0.960731
          Consensus quality: 117571 bases at least Q40
          Consensus quality: 145749 bases at least Q30
          Consensus quality: 160940 bases at least Q20
          Insert size: 185000; agarose-fp
          Insert size: 171264; sum-of-ctigs
          Quality coverage: 2.9 in Q20 bases; agarose-fp
          Quality coverage: 3.2 in Q20 bases; sum-of-ctigs
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 39 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          * 1003: contig of 1003 bp in length
          * 1004 1103: gap of 100 bp
          * 1104 2634: contig of 1531 bp in length
          * 2635 2734: gap of 100 bp
          * 2735 4415: contig of 1681 bp in length
          * 4416 4515: gap of 100 bp
          * 4516 5785: contig of 1270 bp in length
          * 5786 5885: gap of 100 bp
          * 5886 7879: contig of 1994 bp in length
          * 7880 7979: gap of 100 bp
          * 7980 9686: contig of 1707 bp in length
          * 9687 9786: gap of 100 bp
          * 9787 12253: contig of 2467 bp in length
          * 12254 12353: gap of 100 bp
          * 12354 15228: contig of 2875 bp in length
          * 15229 15328: gap of 100 bp
          * 15329 17200: contig of 1872 bp in length
          * 17201 17300: gap of 100 bp
          * 17301 20131: contig of 2831 bp in length
          * 20132 20231: gap of 100 bp
          * 20232 22587: contig of 2356 bp in length
          * 22588 22687: gap of 100 bp
          * 22688 25707: contig of 3020 bp in length
          * 25708 25807: gap of 100 bp
          * 25808 28184: contig of 2377 bp in length
          * 28185 28284: gap of 100 bp
          * 28285 31338: contig of 3054 bp in length
          * 31339 31438: gap of 100 bp
          * 31439 34299: contig of 2861 bp in length
          * 34300 34399: gap of 100 bp
          * 34400 38318: contig of 3919 bp in length
          * 38319 38418: gap of 100 bp
          * 38419 42835: contig of 4417 bp in length
          * 42836 42935: gap of 100 bp
          * 42936 45448: contig of 2513 bp in length
          * 45449 45548: gap of 100 bp
          * 45549 48116: contig of 2568 bp in length
          * 48117 48216: gap of 100 bp
          * 48217 52618: contig of 4402 bp in length
          * 52619 52718: gap of 100 bp
          * 52719 56592: contig of 3874 bp in length
          * 56593 56692: gap of 100 bp
          * 56693 59635: contig of 2943 bp in length
          * 59636 59735: gap of 100 bp
          * 59736 63661: contig of 3926 bp in length
          * 63662 63761: gap of 100 bp
          * 63762 68437: contig of 4676 bp in length
          * 68438 68537: gap of 100 bp
          * 68538 71458: contig of 2921 bp in length
          * 71459 71558: gap of 100 bp
          * 71559 76888: contig of 5330 bp in length
          * 76889 76988: gap of 100 bp
          * 76989 82113: contig of 5125 bp in length
          * 82114 82213: gap of 100 bp
          * 82214 88220: contig of 6007 bp in length
          * 88221 88320: gap of 100 bp
          * 88321 93499: contig of 5179 bp in length
          * 93500 93599: gap of 100 bp
          * 93600 97901: contig of 4302 bp in length
          * 97902 98001: gap of 100 bp
          * 98002 103016: contig of 5015 bp in length
          * 103017 103116: gap of 100 bp
          * 103117 109178: contig of 6062 bp in length
          * 109179 109278: gap of 100 bp
          * 109279 117307: contig of 8029 bp in length
          * 117308 117407: gap of 100 bp
          * 117408 124079: contig of 6672 bp in length
          * 124080 124179: gap of 100 bp
          * 124180 131281: contig of 7102 bp in length
          * 131282 131381: gap of 100 bp
          * 131382 138059: contig of 6678 bp in length
          * 138060 138159: gap of 100 bp

```

	Query Match	44.5%;	Score 159;	DB 2;	Length 175064;
	Best Local Similarity	100.0%;	Pred. No. 2.7e-81;		
	Matches 159;	Conservative	0;	Mismatches	0; Indels 0; Gaps
QY	199	tgttgccctcagctgaggtgtgctgctgtggaagaacctcacttcagaagaagacaaaca	258		
Db	151314	TGTTGGCCTCAGCTGAGGTGTGCTGTGTGGAAGAACTTCACCTTCAGAAGAAGACAAACA	151255		

21251004
MEDLINE
REFERENCE
2 (bases 1 to 201144)
AUTHORS
Qiu, Y., Cavellier L., Chiu, S., Rubin, E. and Cheng, J.-F.
TITLE
Direct Submission
JOURNAL
Submitted (13-JUL-2000) Genome Science Department, Lawrence

Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley, CA 94720, USA

FEATURES

Location/Qualifiers
1. 201144

1. 207144

```
location/qualifiers
1. .201144
/organism="Homo sapiens"
```

mRNA

/sub_xref=taxon:9606"
 join(33931..34151,58319..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,122329..122569,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,139268..139381,140379..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150428,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170229,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..181457)/
 /gene="ABCA1"
 /product="ATP-binding cassette 1 sub-family A member 1"
 33931..181457
 /gene="ABCA1"
 33931..34151
 /gene="ABCA1"
 /number=1
 58319..58476
 /gene="ABCA1"
 /number=2
 join(58411..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,122329..122569,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150428,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170229,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..178700)/
 /gene="ABCA1"

product="Ap-
plication/

```

/cocon_s_cat=1
/product="ATP-binding cassette 1 sub-family A member 1"
/protein_id="AA042526.1"

```

```
/procII_ID=AAK43526:  
/db_xref="GI:13876613"
```

```
/protein_id=AAK43526.1"
/db xref=GI:13876613"
```

[illegible]

STVSYLKKEDSVSQSSDAGLSHSDHSDILITDVSAISNIRKHVSEARLVEDIGHET
 TVYPLERAKGAFVFLHEIDRDLSDLGSSVGSSETTLEEIFLKVAEESGVDAETS
 DGLTPARRNRAPFGDQKSLRPTEDDADPNDSIDPESRETDLLSGMDGKGSYQVK
 GWLTKQOQFVALLWKKLLIARRSKGFFAIYILVPAPFVFCIALVFSLIVPPFGKYPSLE
 LKPMYNEQVTVFSNDAPEDTGTLELLNLTKDQPGFTRCMEGNPIDTPCOAGBREW
 TTAAPQTTMDLFPNGNWTMNPSPACQSSDKIKMLPVCPPGAGGLPPQOKONTA
 DILQDITGRNISDLYLVTVQIIIAKSLKNKIVNERRYGGSILGVSNTOALPPSQBVN
 DAILOKMKHLKAKDSSADRFNLNSLGRFWTGLDTRNNVAVWNNKGWHAISSELNVIN
 NALLRANLQERNPSHYGITAFNHPNLNTKQOLSEVALMTTSDVLVSTCVTFAMSFV
 PASFVNLQIQRVSKAKHLQFISGVKPVYIWLNSVMDMCMYVVPATLVIIFICFQO
 KSVSTNEPLVIALALLXGWSITPLMYPASVFKTIPSTAYVVLTSVNLFIGINGSVA
 TVLEFLTQNKLNININDILKSVFLIFPHFLCGLGLDMVKNOAMADALERFGENREVS
 PLSLDWGNLNFAMAVEGVVFFLITVLIOYRFIERPVPYNAKLSPLNDEDEDVRRERQ
 RILDGSGNDILEIKELTKIYRKRKPAVDRLCVGIPGCEGFLGNGAGAKSSFTFKM
 LFGDTIVTRGDAPFLNKNLSILNIEHVQNMGTCPQDFDAITELTGREHVEFFALLRGV
 PEKVGKGEWAIRKLGVLKYGEKYAGNYSGGNKRLKSTAMALIGGPPVYFDETTG
 MPOKARFLMNCALSVYKEGRSVLTSHSMCEALCTMAIIVAGRFRLCGLSGVOHLK
 NFRGQDQTTIVRIASGNPDLKPVDQDFGLAFPGSVLKEKHNMLQVLPSSLSGLIARI
 FSTLSQSKRRLHIEDYSVSQTTLDQVFNFAKQSDDDHDKDLSLHKNTVYVDVAVLT
 SFLQDSKVESVY"
 72893. .72986
 /gene="ABCA1"
 /number=3
 77520. .77661
 /gene="ABCA1"
 /number=4
 78931. 79049
 /gene="ABCA1"
 /number=5
 101048. .101169
 /gene="ABCA1"
 /number=6
 104152. .104328
 /gene="ABCA1"
 /number=7
 117287. .117379
 /gene="ABCA1"
 /number=8
 123239. .123269
 /gene="ABCA1"
 /number=9
 125285. .125424
 /gene="ABCA1"
 /number=10
 125754. .125870
 /gene="ABCA1"
 /number=11
 130079. 130276
 /gene="ABCA1"
 /number=12
 131023. .131228
 /gene="ABCA1"
 /number=13
 131749. 131925
 /gene="ABCA1"
 /number=14
 133708. .133930
 /gene="ABCA1"
 /number=15
 135677. .135898
 /gene="ABCA1"
 /number=16
 136959. .137163
 /gene="ABCA1"
 /number=17
 138268. .138381
 /gene="ABCA1"
 /number=18
 140179. .140350
 /gene="ABCA1"
 /number=19
 141340. .141471
 /gene="ABCA1"

16778	17678: contig of 901 bp in length
17679	17778: gap of 100 bp
17779	18679: contig of 901 bp in length
18680	18779: gap of 100 bp
18780	19632: contig of 853 bp in length
19633	19732: gap of 100 bp
19733	20634: contig of 902 bp in length
20635	20734: gap of 100 bp
20735	21620: contig of 886 bp in length
21621	21720: gap of 100 bp
21721	22579: contig of 859 bp in length
22580	22679: gap of 100 bp
22680	23568: contig of 889 bp in length
23569	23668: gap of 100 bp
23669	24554: contig of 886 bp in length
24555	24654: gap of 100 bp
24655	25521: contig of 867 bp in length
25522	25621: gap of 100 bp
25623	26487: contig of 866 bp in length
26488	26587: gap of 100 bp
26588	27464: contig of 877 bp in length
27465	27564: gap of 100 bp
27565	28466: contig of 902 bp in length
28467	28566: gap of 100 bp
28567	29464: contig of 898 bp in length
29465	29564: gap of 100 bp
29565	30447: contig of 883 bp in length
30448	30547: gap of 100 bp
30548	31453: contig of 906 bp in length
31454	31553: gap of 100 bp
31554	32452: contig of 899 bp in length
32453	32552: gap of 100 bp
32553	33447: contig of 895 bp in length
33448	33547: gap of 100 bp
33548	34435: contig of 888 bp in length
34436	34535: gap of 100 bp
34536	35433: contig of 898 bp in length
35434	35533: gap of 100 bp
35534	36440: contig of 907 bp in length
36441	36540: gap of 100 bp
36541	37422: contig of 882 bp in length
37423	37522: gap of 100 bp
37523	38402: contig of 880 bp in length
38403	38502: gap of 100 bp
38503	39380: contig of 878 bp in length
39381	39480: gap of 100 bp
39481	40407: contig of 927 bp in length
40408	40507: gap of 100 bp
40508	41405: contig of 898 bp in length
41406	41505: gap of 100 bp
41506	42399: contig of 894 bp in length
42400	42499: gap of 100 bp
42500	43357: contig of 858 bp in length
43358	43457: gap of 100 bp
43458	44356: contig of 859 bp in length
44357	44456: gap of 100 bp
44457	45325: contig of 869 bp in length
45326	45425: gap of 100 bp
45426	46305: contig of 880 bp in length
46306	46405: gap of 100 bp
46406	47302: contig of 897 bp in length
47303	47402: gap of 100 bp
47403	48293: contig of 891 bp in length
48294	48393: gap of 100 bp
48394	49280: contig of 887 bp in length
49281	49380: gap of 100 bp
49381	50250: contig of 870 bp in length
50251	50350: gap of 100 bp
50351	51209: contig of 859 bp in length
51210	51309: gap of 100 bp
51310	52221: contig of 912 bp in length
52222	52321: gap of 100 bp
52322	53205: contig of 884 bp in length

[illegible]

```

FEATURES
  source
    Location/Qualifiers
      1..99
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT
  17 a 22 c 27 g 33 t
ORIGIN

```



```

RESULT 14
AX060892
LOCUS
DEFINITION Sequence 1 from Patent WO0078971.
ACCESSION AX060892
VERSION AX060892.1 GI:12406270
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE ATP binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Source
1..10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
ORIGIN
Query Match 21.6%; Score 77; DB 6; Length 10442;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggcgtgctctccagggtcccgagcacacgctggcgctgctgagga 190
Db 229 AGCTTGCGCGCTGCTTCCAGGGCTCCGAGGACACGCTGGCGTGGCTGAGGA 288
QY 191 acatggcatgttgctcagctgaggttgcgtgtggaagaacctcacttcagaaga 250
Db 289 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGCTGTGGAAGAAGCTCCTTTCAGAA 348
QY 251 gacaaaca 258
Db 349 GACAAACA 356
RESULT 15
AF285167
LOCUS
DEFINITION Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA,
complete cds.
ACCESSION AF285167
VERSION AF285167.1 GI:9755158
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Schwartz,K., Lawn,R.M. and Wade,D.P.
TITLE ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
regulated by LXR
JOURNAL Unpublished
AUTHORS Lawn,R.M., Wade,D.P., Garvin,M.R., Wang,X., Schwartz,K.,
Porter,J.G., Sellmeyer,J.J., Vaughan,A.M. and Oram,J.F.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-2000) Discovery Research, CV Therapeutics Inc.,
3172 Porter Drive, Palo Alto, CA 94304, USA
FEATURES
Source
1..10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/cell_type="fibroblast"
/tissue_type="skin"
1..10442

```

CDS

```

/gene="ABCA1"
291..7076
/gene="ABCA1"
/codon_start=1
/product="ATP-binding cassette transporter 1"
/protein_id="AAF98175.1"
/db_xref="GI:9755158"
/translacion="MACWPQLRLLLMKNTLFRRTQCQLLLEAVMPLFLLILISVRL
SYFPOHECHFFNKASGTLPPVQGIICANNPCFRPYTPEAGVGVNFKSIV
ARLFSARLLLYSOKDMSKMRKVLQTLQIKSSNNLKQDFLDVNETSGFLYH
NLSPKSTVDKMLRADVILHVKVQYQLHLTKSSNGSEEMIOIGDQVSELCGLP
KEKLAARVLRNMDILKPLILNSTSPSPSKELAEATKTLHSLGTLAQELFSMR
SWSDMRQEWFLTNVSSSTOIYQASRIVCVCHGEGGKIKSLNLTEDNNYKALF
GGNGTEADAEFTYDNTSTPYCNLDMKNLESPSLIIRKALKPLLVKGLTYLPTPAT
ROYMAEVNKTQELAVHLEGGMEELSPKIMTFMENSQEMDLRLSLRNDHNEWLN
QQLDGLTQADQIAVFAKHPEDVQSSNGSVTWFRAFNETQAIIRSFNECVNKL
KLEPIATEVWLINKMELLDKRFAGIVFTGTPGSIIEPHHVYKIKMDINDMPT
NKIKDGYWDPGPRADPFEDMRVWVGFAVLDQVVEQAIIRLVTHGKTKGVYQOOPY
PCVYDDIFLRVMSRSMPLFWTLAWIYSVAIIIGIVYEKARKETMRIMGLDNLILW
FSWFISSLIPLLVSAGLLVILKGNLLPYSDPVVFLSVFVAVVILQCFLLSTLF
SRANLAACGGIIYFTLYLPYVICVAVQDYVGFTLKIFASLSPVAFGCGCYFALFE
EGGIGVQNDLFESEDEGPNLTTSISMLEDFTLYGYMTWYIEAFPGQYIGIPRW
YFPTKSYWFEESDEKSHPGSNQKRMSEICMEEPHLKGLVSIQNLVKKVYDGMKV
AVDGLALNFEQITSLFNGHAGKTTMSILTLGLPPTSGTAYILGDIRSEMTIR
QNLGVCPOHNVLDMTLVBEHTWFAVLKGLSEKHKVKAEMQALDGLPSSKLKST
SOLSGMQKLSVALAFVGSKVILDEPTAGVDYVSRGIWELLKLLKQGRITILST
HMDAEDVLGDRALISHGKLCVGSLLKQNLGTYLTLVKKDSESSLCRNS
SVYSYLKEDSSVQSSDAGLDHSDTLTIDVSAISNLIRKHSEARLVEDIGHEL
TLYPYEAKRGAFAVELFHEIDRLSDGISYSETLEIFLLKABESGVDAETS
DGTLPARRNRAGFKQSCRLPTEDDAADPDSIDPESRETFLLSGMDGSGQVK
GWLTKQQFVALLKRLIARRSRGFAIVLPAFVFCIALVSLIYVPPFKYPSLE
LQPMYNEQYTFVSNDAEDTGTLELLNALTDPGFCRMEGNLIPDTQCAEGEEM
TTAPVPTIMDLFONGMTQNPSPACQSSDKIKMLPVCPPGAGGLPPQKOKNTA
DILODGTENISDYLVKTVQIILAKSLNKHINWNEFRYGGFSLGVSNQALPPSOEVN
DAIKOMKHLKAKLSSADRLNSLGRMTGLDRNNKWNFNKNGHAISSFLVIN
NAILRANLOKGENSHYGFATFNHPLNTKQSLSEVALMTSDVLYVICVIFAMSFY
PASVYFLIQRVSKAKHLQFISGVKPVYIWLSEFVMDMNTVVPATLVIIIFCQO
KSYVSTNLPVIALLLLYGWSITPLMPYPAFVKIPSTAYVLTSLVNLFTINGSA
TFVLELTDKNLNNIDILKSVLFIHPHCLGRGLIDMKVQKADALERGENRFS
PLSVDLVRNLFAVAGEVVFELITVLOVREFIIRPVNAKLSPLNDEDEVRERQ
RILDGGQNDLLEIKELIKIYRKRKPAVDRCVIGIPGECGLLGVNGAGKSTFKM
LIGDTVTIRGDALFNKNSILSNHVNQNYCPOQDAITELLTGREHVEFFALLRGV
PEKEVGGEWAIRKGLVYKGYAGNYSGNKRKLSTAMALIGGPVFLDEPTTG
MDPKARRLWNCALSVKRGSVLTSMECEALCTRMALIMVNGRRCVGLSGVOHLK
NRFSGDITVIRIAGSNPDLKPVQDFGLAFPGSVLKEKHNMLQYQLPSSLSIARI
FSLGSQSKRLHIEDYSVSQITLDQVVFVNEAKDQSDDDHLKLSLHKNQTVYDAVLT
SFLQDEKVKESVY"
BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
ORIGIN

```

```

Query Match 21.6%; Score 77; DB 9; Length 10442;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggcgtgctctccagggtcccgagcacacgctggcgctgagga 190
Db 229 AGCTTGCGCGCTGCTTCCAGGGCTCCGAGGACACGCTGGCGTGGCTGAGGA 288
QY 191 acatggcatgttgctcagctgaggttgcgtgtggaagaacctcacttcagaaga 250
Db 289 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGCTGTGGAAGAAGCTCCTTTCAGAA 348
QY 251 gacaaaca 258
Db 349 GACAAACA 356

```

```

RESULT 16
AX060719
LOCUS
DEFINITION Sequence 7 from Patent WO0078972.
ACCESSION AX060719

```

linear PAT 22-JAN-2001

```

VERSION      AX060719.1  GI:12406108
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE       1 (bases 1 to 10474)
JOURNAL     Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
            Regulation with binding cassette transporter protein abcl
            Patent: WO 0078972-A 7 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES
SOURCE      1..10474
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  2906 a 2305 c 2416 g 2843 t 4 others
ORIGIN
Query Match      21.6%; Score 77; DB 6; Length 10474;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctggcgagggga 190
      |||||||
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCGAGCCACACGCTGGGGTGTGCTGGTGGGGA 320
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 250
      |||||||
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGTTGGAAGAACCTCACTTTCAGAAGAA 380
QY 251 gacaaaca 258
      |||||||
Db 381 GACAAACA 388

RESULT 17
AX060721
LOCUS      AX060721 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078972.
ACCESSION  AX060721
VERSION    AX060721.1  GI:12406109
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE     1 (bases 1 to 10474)
JOURNAL   Lawn,R.M., Wade,D. and Garvin,M.
            Regulation with binding cassette transporter protein abcl
            Patent: WO 0078972-A 9 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES
SOURCE      1..10474
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN
Query Match      21.6%; Score 77; DB 6; Length 10474;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctggcgagggga 190
      |||||||
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCGAGCCACACGCTGGGGTGTGCTGGTGGGGA 320
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 250
      |||||||
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGTTGGAAGAACCTCACTTTCAGAAGAA 380
QY 251 gacaaaca 258
      |||||||
Db 381 GACAAACA 388

RESULT 18
AX060898
LOCUS      AX060898 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 7 from Patent WO0078971.
ACCESSION  AX060898
VERSION    AX060898.1  GI:12406275
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE     1 (bases 1 to 10474)
JOURNAL   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
            Atp binding cassette transporter protein abcl polypeptides
            Patent: WO 0078971-A 7 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES
SOURCE      1..10474
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  2906 a 2305 c 2416 g 2843 t 4 others
ORIGIN
Query Match      21.6%; Score 77; DB 6; Length 10474;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctggcgagggga 190
      |||||||
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCGAGCCACACGCTGGGGTGTGCTGGTGGGGA 320
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 250
      |||||||
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGTTGGAAGAACCTCACTTTCAGAAGAA 380
QY 251 gacaaaca 258
      |||||||
Db 381 GACAAACA 388

RESULT 19
AX060900
LOCUS      AX060900 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078971.
ACCESSION  AX060900
VERSION    AX060900.1  GI:12406276
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE     1 (bases 1 to 10474)
JOURNAL   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
            Atp binding cassette transporter protein abcl polypeptides
            Patent: WO 0078971-A 9 28-DEC-2000;
            CV THERAPEUTICS, INC. (US)
FEATURES
SOURCE      1..10474
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN
Query Match      21.6%; Score 77; DB 6; Length 10474;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctggcgagggga 190
      |||||||
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCGAGCCACACGCTGGGGTGTGCTGGTGGGGA 320
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaa 250
      |||||||
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGTTGGAAGAACCTCACTTTCAGAAGAA 380
QY 251 gacaaaca 258
      |||||||
Db 381 GACAAACA 258
```

KEYWORDS	synthetic construct.									
SOURCE	synthetic construct									
ORGANISM	artificial sequence.									
REFERENCE	1 (bases 1 to 446)									
AUTHORS	Denefle,P., Rosier-Montus,M.F., Arnould-Requigne,I., Prades,C., Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H., Remaley,A., Brewer,H.B. and Dean,M.									
TITLE	Nucleic acids of the human abcl gene and their therapeutic and diagnostic application									
JOURNAL	Patent: WO 0130848-A 3 03-MAY-2001;									
FEATURES	Aventis Pharma S.A. (FR)									
source	Location/Qualifiers									
	1..446									
	/organism="synthetic construct"									
	/db_xref="taxon:32630"									
BASE COUNT	96 a 123 c 112 g 115 t									
ORIGIN										
	Query Match 16.8%; Score 60; DB 6; Length 446;									
	Best Local Similarity 100.0%; Pred. No. 1.7e-23;									
	Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps									
QY	199 tgttgccctcagctgaggtgctgctgtggaagaccctcacttcagagaagacaaaca 258									
Db	191 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAGAACCTCAGTTTCAGAGAAGACAAACA 250									
RESULT 22										
LOCUS	AX139751									
DEFINITION	Sequence 3 from Patent EP1096012.									
ACCESSION	AX139751									
VERSION	AX139751.1 GI:14275333									
KEYWORDS	synthetic construct.									
SOURCE	synthetic construct									
ORGANISM	artificial sequence.									
REFERENCE	1 (bases 1 to 446)									
AUTHORS	Denefle,P., Rosier-Montus,M.F., Arnould-Requigne,I., Prades,C., Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,Ill,G.H., Remaley,A., Brewer,H.B. and Dean,M.									
TITLE	Nucleic acids of the human abcl gene and their therapeutic and diagnostic application									
JOURNAL	Patent: EP 1096012-A 3 02-MAY-2001;									
FEATURES	Aventis Pharma S.A. (FR)									
source	Location/Qualifiers									
	1..446									
	/organism="synthetic construct"									
	/db_xref="taxon:32630"									
BASE COUNT	96 a 123 c 112 g 115 t									
ORIGIN										
	Query Match 16.8%; Score 60; DB 6; Length 446;									
	Best Local Similarity 100.0%; Pred. No. 1.7e-23;									
	Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps									
QY	199 tgttgccctcagctgaggtgctgctgtggaagaccctcacttcagagaagacaaaca 258									
Db	191 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAGAACCTCAGTTTCAGAGAAGACAAACA 250									
RESULT 23										
LOCUS	AF258627									
DEFINITION	Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,									
ACCESSION	AF258627									
VERSION	AF258627.1 GI:7769707									
LOCUS	AF258627									
DEFINITION	Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,									
ACCESSION	AF258627									
VERSION	AF258627.1 GI:7769707									

KEYWORDS SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 697)
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J., and Kane, J.P.
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 1 to 697)
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J., and Kane, J.P.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
FEATURES Location/Qualifiers
source 1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9c31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/note="ABC1"
396..>697
/gene="ABCA1"
/note="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AAF69513.1"
/db_xref="GI:7769708"
/translation="MACWPOLRLKLNLFRRROTCOLLLEVAWPLFIFLILISVRL
SYPYEHECHFPKAMPSAGTLPWVGIIICNANPCFRPTGPEAGPVVGNFKS
152 a 198 c 190 g 156 t 1 others
BASE COUNT
ORIGIN
Query Match 16.8%; Score 60; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 ttttgccctcagctgaggttctgtctgtggaagaacctcactttcagaagaagacaaca 258
|||||
Db 402 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACA 461
RESULT 24
AB055982
LOCUS Homo sapiens mRNA for ABCA1, complete cds. PRI 18-AUG-2001
DEFINITION
ACCESSION AB055982
VERSION AB055982.1 GI:15212106
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6786)
Tanaka, A.R., Abe-Dohmae, S., Arakawa, R., Sadanami, K., Kidera, A.,
Kioka, N., Amachi, T., Yokoyama, S. and Ueda, K.
TITLE A new topological model of functional human ABCA1-Signal peptide
cleavage and glycosylation of a large extracellular domain
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6786)
Ueda, K., Kioka, N. and Tanaka, A.R.
AUTHORS Direct Submission
TITLE Submitted (20-FEB-2001) Kazumitsu Ueda, Kyoto University Graduate
School, Applied Life Sciences, Kitashirakawa, Kyoto Sakyo-ku, Kyoto
606-8502, Japan (E-mail:uedak@kais.kyoto-u.ac.jp).

Tel:81-75-753-6105, Fax:81-75-753-6104)
FEATURES source
Location/Qualifiers
1..6786
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..6786
/gene="ABCA1"
1..6786
/gene="ABCA1"
/codon_start=1
/product="ABCA1"
/protein_id="BAB63210.1"
/db_xref="GI:15212107"
/translation="MACWPOLRLKLNLFRRROTCOLLLEVAWPLFIFLILISVRL
SYPYEHECHFPKAMPSAGTLPWVGIIICNANPCFRPTGPEAGPVVGNFKS
ARLFSADARLLIYKODISGMKMRVLRLOIKKSSNLKLODLVDFNVEFSGFLYH
NLSLPKSTVDKMLRMDVILHKVLOGYQLHLSLCLNGSEMIQGLQGVSELCGLP
RKLAAERVLRSMDILKPIRLTNSPFPKELAEATKTLHSLGLTLAGELSMR
SWSMDRQEVFLTNVSSSSQIYQAVSRIVCGHPEGGLKIKSLNWEEDNYYKALF
GNGTEEDAETFDNSTTTCNDLMKNLSSPLRIWKALPLLVGKTLIYPTDPAT
QVMAEVNKTFOELAVFHDLEGWELSPKWTMENSOEMLVRLMLDSRDHCVNLN
QDLGLDWTQADIVAFIAKHPEDVQSSNGSVTWREAFNETQOALRTISREWCYNL
NKIKDGYNDPGRADPDMRYWVGGFAYLQDVEQAILRVLTGTEKKRGVYQWQMPY
PCYVDDIFLRVMSRMPLEMTLAWIYSAVILKIGVKEARLKETMRIMLNDLSILM
FSWFTSLIPLVSLAGLVVILKLNLPYSDPSVVFVLSVFAVVTILQCELIETLF
SRANLAAACGGIIFYTLIYVLCVADQYVGTFLKIFASLLSPVAFGCGCEVFALE
EQGIGVQNDNLFESPVEDGFNLITTSVSMMLPDTLYGYMTWYIEAVFGPGYIGRPW
YFPTCKSYWFEESDEKSHPGSKRRIKSEIEMEEETHLKLGVSQNLVYKVRDGMKV
AVDGLANFYEGQITSLFHNGAGKTTMSILTLGLPPTSGTAYILGKDIRSEMSTIR
QNLGVCQHNVLDFMLTBEHIFVARKLSEKHKVMEQMAQLDGLPSSKLSKT
SOLSGMQRLKVALAFVGGSKVYILDEPTAGVDPYSRRGIWELLIKYRQRTIILST
HMDADVLGDRATISHGKLCVGSLLFKNLQGTGYTLVKKVSEARLVEDIGHEL
STVYPLAEKAGFAELFHEIDRLSDIGISYSGSETLEELFLKVAESGVDAETS
DGTLPARRNRAPFGQCSCLPTEDDADNDSDIDPESRETLDSGMDGKSTQVK
GKLTQQQVALLKRLLIARRSKGFFAIPLPAVFVCIALFSLIYVPPFKYPSLE
LQPMYNEQTVFVSDAPEDTGTLELNLATKDPGFCRCMEGNPILPDPQAGEEEM
TTAPVQITMDLPQNGWTMPNSPACQSSDKIKMLPVCPPGAGGLPQALPPSQNTA
DILQDLTGRNISDLYLVYQILAKSLKNKIVWNEPRYGGFSLGVSNQALPPSQBN
DAIKQMKHLKAKSSADRELSLGRFMTGLDTKNKVKWNNKGNHAISSFLANVIN
NAILRANLQGNPSHYGITAFNHLNLTQKQSEVALMTTSDVLVLSICVAFMSFV
PASVYFLIQLERYSKAKHLQFSGVPRYIYWSNFVCMYVVPATLVIIIFICFQQ
KSVSTNLPVALLKLLQWSTPLMYPASFVKIPSTAYVYVLTNLFINGINGVA
TFVLELTNKLNNIINDILKSVLELPHFCLGRLDMVKNQAMADALERENREVS
PLSMDLVGRNLFAMAVEGVVFLITVLIQYRFRPRPNKLSLNGDEDEDVRRRQ
RILDGQNDLLEIKELTKYRBRKPAVDICVGTIPGCGCLLGVNGAGKSSFTKM
LTGDTTVTRGDAFLNKNLSLNIHEVQNMGCYQFDATITELLTGREHVEFALLRGV
PEKEVGKGEWAIRKIGLVYKGYAGNYSGGNKKRLSTAMALIGPPVYVFLDEPTTG
MDPKARFLWNCALSVYKGRSVLTSMSHECEALCTRNAIMVNGFRCLGSGVQHLK
NRFQDGYTIWRIAGSNPDLKPVQDFGLPFGSVLKEKRRNMLOQLPSSLSLARI
FSILSOSKRLRIEDYSVSTLTLDQVFNFAKQDSDHDLKDLSLHKNQTVVDVAULT
SFLQDKVRESYV"
BASE COUNT 1724 a 1643 c 1759 g 1660 t
ORIGIN
Query Match 16.8%; Score 60; DB 9; Length 6786;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 ttttgccctcagctgaggttctgtctgtggaagaacctcactttcagaagaagacaaca 258
|||||
Db 7 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAACA 66
RESULT 25
AX253452
LOCUS Sequence 3 from Patent WO0170810.
DEFINITION
ACCESSION AX253452
VERSION AX253452.1 GI:16073979
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE 1 (bases 1 to 7260)
Schmitz, G. and Bodzioch, M.
JOURNAL ATP binding cassette transporter 1 (abcl) gene polymorphisms and
cardiovascular or inflammatory disorders
Patent: WO 0170810-A 3 27-SEP-2001;
Bayer Aktiengesellschaft (DE)
FEATURES Location/Qualifiers
source 1..7260
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1834 a 1765 c 1905 g 1756 t
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 7260;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 tgttgccctcagctgaggttgcctgtggaagaacccctcactttcagaagaagacaaca 258
|||||
Db 327 TGTGGCCTCAGCTGAGGTGCTGCTGGAGAACCCTCAGCTTCAGAGAAGACAACA 386
|||||
RESULT 26
LOCUS AX092594 7860 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 6 from Patent WO0115676.
ACCESSION AX092594
VERSION AX092594.1 GI:13444651
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE 1 (bases 1 to 7860)
Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.
JOURNAL Compositions and methods for modulating hdl cholesterol and
triglyceride levels
Patent: WO 0115676-A 6 08-MAR-2001;
University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES Location/Qualifiers
source 1..7860
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2014 a 1860 c 2008 g 1978 t
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 7860;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 tgttgccctcagctgaggttgcctgtggaagaacccctcactttcagaagaagacaaca 258
|||||
Db 81 TGTGGCCTCAGCTGAGGTGCTGCTGGAGAACCCTCAGCTTCAGAGAAGACAACA 140
|||||
RESULT 27
LOCUS AX135712 7862 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 1 from Patent WO0132184.
ACCESSION AX135712
VERSION AX135712.1 GI:14271961
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 7862)
AUTHORS Attie, A.D., Cook, M., Gray-Keller, M.P., Hayden, M.R., Pimstone, S. and
Brooks-Wilson, A.
TITLE Abcl modulation for the modulation of cholesterol transport
JOURNAL Patent: WO 0132184-A 1 10-MAY-2001;
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES Location/Qualifiers
source 1..7862
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2013 a 1861 c 2010 g 1978 t
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 7862;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 tgttgccctcagctgaggttgcctgtggaagaacccctcactttcagaagaagacaaca 258
|||||
Db 81 TGTGGCCTCAGCTGAGGTGCTGCTGGAGAACCCTCAGCTTCAGAGAAGACAACA 140
|||||
RESULT 28
LOCUS AX127830 9741 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 69 from Patent WO0130848.
ACCESSION AX127830
VERSION AX127830.1 GI:14134477
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE 1 (bases 1 to 9741)
Deneffe, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,
Remaley, A., Lemoine, C., Duverger, N., Jaye, M., Searfoss, G.H.,
Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
Patent: WO 0130848-A 69 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
source 1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 tgttgccctcagctgaggttgcctgtggaagaacccctcactttcagaagaagacaaca 258
|||||
Db 191 TGTGGCCTCAGCTGAGGTGCTGCTGGAGAACCCTCAGCTTCAGAGAAGACAACA 250
|||||
RESULT 29
LOCUS AX139817 9741 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 69 from Patent EP1096012.
ACCESSION AX139817
VERSION AX139817.1 GI:14275399
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE 1 (bases 1 to 9741)
Deneffe, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,
Naudin, L., Lemoine, C., Duverger, N., Jaye, M., Searfoss, G.H.,

Remaley, A., Brewer, H.B. and Dean, M.
Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 69 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
source 1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 ttttgccctcagctgaggttgcctgtgtggaagaacctcactttcagaagaagacaaca 258
Db 191 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCCTTTTCAGAGAAGACAACA 250
RESULT 30
AX127831
LOCUS 9741 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 10 from Patent WO0183746.
ACCESSION AX351038
VERSION AX351038.1 GI:18616393
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9741)
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P.,
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 10 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
source 1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 ttttgccctcagctgaggttgcctgtgtggaagaacctcactttcagaagaagacaaca 258
Db 191 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCCTTTTCAGAGAAGACAACA 250
RESULT 31
AX127831
LOCUS 9854 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 70 from Patent WO0130848.
ACCESSION AX127831
VERSION AX127831.1 GI:14134478
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9854)
AUTHORS Denefle, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,
Naudin, L., Lemoine, C., Duverger, N., Jaye, M., Searfoss, G.H.,
Remaley, A., Brewer, H.B. and Dean, M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application

JOURNAL Patent: WO 0130848-A 70 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
source 1..9854
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2665 a 2219 c 2334 g 2635 t 1 others
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 9854;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 ttttgccctcagctgaggttgcctgtgtggaagaacctcactttcagaagaagacaaca 258
Db 304 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCCTTTTCAGAGAAGACAACA 363
RESULT 32
AX139818
LOCUS 9854 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 70 from Patent EP1096012.
ACCESSION AX139818
VERSION AX139818.1 GI:14275400
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9854)
AUTHORS Denefle, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,
Naudin, L., Lemoine, C., Duverger, N., Jaye, M., Searfoss, Iii, G.H.,
Remaley, A., Brewer, H.B. and Dean, M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 70 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
source 1..9854
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2665 a 2219 c 2334 g 2635 t 1 others
ORIGIN
Query Match 16.8%; Score 60; DB 6; Length 9854;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 ttttgccctcagctgaggttgcctgtgtggaagaacctcactttcagaagaagacaaca 258
Db 304 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCCTTTTCAGAGAAGACAACA 363
RESULT 33
AK024328
LOCUS 1556 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA FLJ14266 fis, clone PLACE1002437, highly similar
to ATP-BINDING CASSETTE TRANSPORTER 1.
ACCESSION AK024328
VERSION AK024328.1 GI:10436685
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens placenta cDNA to mRNA, clone PLACE1
clone: PLACE1002437.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,
Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K.,
Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,

Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Oshima, A.
 NEDO human cDNA sequencing project
 Unpublished (2000)
 2 (bases 1 to 1556)
 Isogai, T. and Otsuki, T.
 Direct Submission
 Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction; 5'- & 3'-end one pass sequencing and clone selection;
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.

FEATURES
 source
 1..1556
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE1002437"
 /tissue_type="placenta"
 /clone_lib="PLACE1"
 /note="cloning vector: pME18FL3"
 314..1405
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB14887.1"
 /db_xref="GI:10436686"
 /translation="MACWPOLRLILKLNLTERRQCQLLELVAMPFLIFILISVRL
 SYPPYEOHECHPEPKMPSAGTLPWQGIICNANNPCRYPTPGAPGVGNFNKSIY
 ARLESDARLLIYSDIYKMDKMRVLRITLQIIRKSSNKLQDFLVNDFSGFLYH
 NLSLPSKTVKMLRADVLILKVFYLGQYQLHLTSLCNGSKSEMIQLGQEVSELGCLP
 KEKLAARVLRNMDILKPIIMDVACDIIAGQLTVPRAAATGDAKPNMMGRET
 LLSICASVPKVFHERHILEHFCVCSVSLEFPKAGIVSFWSAFRIWLKAVFWQ
 HGESMAVWEGQLGLNIAFEVETSIDVG"
 BASE COUNT 380 a 363 c 399 g 414 t
 ORIGIN

Query Match 14.3%; Score 51; DB 9; Length 1556;
 Best Local Similarity 100.0%; Pred. No. 3.3e-18;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 208 cagctagattctactgtggaacacctcacttccttcagaagaagacaaca 258
 |||||
 Db 329 CAGCTGAGGTGCTGCTGGGAAGAACCTCACTTTTCAGAGAAGACAACA 379

RESULT 34
 AC092843
 LOCUS 37 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 255 from Patent WO0115676.
 AC092843
 VERSION AX092843.1 GI:13444900
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 37)
 AUTHORS Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.
 TITLE Compositions and methods for modulating hdl cholesterol and triglyceride levels
 JOURNAL Patent: WO 0115676-A 255 08-MAR-2001;
 University of British Columbia (CA); Xenon Genetics Inc. (CA)
 FEATURES
 source
 1..37
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic Primer"
 4 a 17 c 11 g 5 t

Query Match 9.8%; Score 35; DB 6; Length 37;
 Best Local Similarity 100.0%; Pred. No. 6.3e-09;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 138 gccgcgtcccttcagggtccgcagccagccacagctg 172
 |||||
 Db 1 GCGCGTGCCTTCAGGGCTCCCGAGCCACACGCTG 35

RESULT 35
 AC092841
 LOCUS 38 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 253 from Patent WO0115676.
 AC092841
 VERSION AX092841.1 GI:13444898
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 38)
 AUTHORS Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.
 TITLE Compositions and methods for modulating hdl cholesterol and triglyceride levels
 JOURNAL Patent: WO 0115676-A 253 08-MAR-2001;
 University of British Columbia (CA); Xenon Genetics Inc. (CA)
 FEATURES
 source
 1..38
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic Primer"
 6 a 15 c 5 g 12 t

Query Match 9.8%; Score 35; DB 6; Length 38;
 Best Local Similarity 100.0%; Pred. No. 6.3e-09;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 78 cccctcgtcttcttccttcagtttaagaccagcc 112
 |||||
 Db 1 CCCCTCTGCTTTATCTTCAGTTAATGACACGCC 35

RESULT 36
 AC091466/c
 LOCUS 207659 bp DNA linear HTG 09-AUG-2001
 DEFINITION Mus musculus clone RP23-353G1, WORKING DRAFT SEQUENCE, 10 unordered pieces.
 AC091466
 VERSION AC091466.2 GI:14336552
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 207659)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 207659)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Boguslavsky, L., Boukhaltier, B., Brown, A.,
 Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
 Collins, S., Collymore, A., Cooke, P., DeAtrellano, K., Dewar, K.,
 Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,
 Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhng, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seanan, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (21-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 11, 2001 this sequence version replaced gi:13702847.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L13324

Center clone name: 353_G.1

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 203563 bases at least Q40

Consensus quality: 205378 bases at least Q30

Consensus quality: 206185 bases at least Q20

Insert size: 205000; agarose-fp

Insert size: 206759; sum-of-contigs

Quality coverage: 9.6 in Q20 bases; agarose-fp

Quality coverage: 9.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 10 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* I 1430: contig of 1430 bp in length

* 1431 1530: gap of 100 bp

* 1531 4004: contig of 2474 bp in length

* 4005 4104: gap of 100 bp

* 4105 132075: contig of 127971 bp in length

* 132076 132175: gap of 100 bp

* 132176 135638: contig of 3463 bp in length

* 135639 135738: gap of 100 bp

* 135739 141621: contig of 5883 bp in length

* 141622 141721: gap of 100 bp

* 141722 157807: contig of 16086 bp in length

* 157808 157907: gap of 100 bp

* 157908 173725: contig of 15818 bp in length

* 173726 173825: gap of 100 bp

* 173826 184364: contig of 10539 bp in length

* 184365 184464: gap of 100 bp

* 184465 206183: contig of 21719 bp in length

* 206184 206283: gap of 100 bp

* 206284 207659: contig of 1376 bp in length.

Location/Qualifiers

1. .207659

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP23-353G1"

/clone_lib="RPCI-23 Female Mouse BAC"

1. .1430

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

misc_feature

1531. 4004

/note="assembly_fragment"

4105. .132075

/note="assembly_fragment"

132176. .135638

/note="assembly_fragment"

135739. .141621

/note="assembly_fragment"

141722. .157807

/note="assembly_fragment"

157908. .173725

/note="assembly_fragment"

misc_feature

4105. .132075

/note="assembly_fragment"

132176. .135638

/note="assembly_fragment"

135739. .141621

/note="assembly_fragment"

141722. .157807

/note="assembly_fragment"

157908. .173725

/note="assembly_fragment"

173826. .184364

/note="assembly_fragment"

184465. .206183

/note="assembly_fragment"

206284. .207659

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT

58917 a 43251 c 44962 g 59609 t 920 others

ORIGIN

Query Match

Best Local Similarity 7.6%; Score 27; DB 2; Length 207659;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88

tttatcttcagtaataatgaccagccac 114

|||||

TTTATCTTCAGTAAATGACCAGCCAC 175436

RESULT 37

AF287263

LOCUS

DEFINITION

Mus musculus ATP-binding cassette 1, sub-family A, member 1 (Abca1)

gene, complete cds.

AF287263.1 GI:11611824

VERSION

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 278572)

Qiu, Y., Cavellier, L., Chiu, S., Yang, X., Rubin, E. and Cheng, J.F.

Human and mouse abca1 comparative sequencing and transgenesis

studies revealing novel regulatory sequences

Genomics 73 (1), 66-76 (2001)

21251004

2 (bases 1 to 278572)

Qiu, Y., Cavellier, L., Chiu, S., Rubin, E. and Cheng, J.F.

Direct Submission

Submitted (14-JUL-2000) Genome Science Department, Lawrence

Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,

CA 94720, USA

Location/Qualifiers

1. .278572

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="BAC 129K10: 142M4; 197F14"

join(87184. .87402,103551.103659,115064. .115187,

119217. .119358,120364. .120482,133479. .133600,

135582. .135758,147083. .147175,149791. .150031,

152367. .152506,152845. .152961,156570. .156767,

157459. .157664,158200. .158376,158821. .159043,

160617. .160838,161790. .161994,163076. .163189,

164646. .164817,165738. .165869,167171. .167313,

167492. .167629,168320. .168540,169790. .169862,

170735. .170937,172636. .172684,172855. .172968,

173925. .174073,175605. .175729,176761. .176859,

FEATURES

source

misc_feature

1531. 4004

/note="assembly_fragment"

4105. .132075

/note="assembly_fragment"

132176. .135638

/note="assembly_fragment"

135739. .141621

/note="assembly_fragment"

```

180181. 180370,181179. 181273,182395. 182427,
183398. 183503,191761. 191829,192306. 192475,
193631. 193808,196619. 196734,196845. 196989,
198526. 198649,200922. 201051,201281. 201401,
202114. 202176,202890. 202996,204955. 205096,
205427. 205558,206739. 206842,207361. 207453,
208135. 208378,209638. 210783)
/gene="Abcal"
/product="ATP-binding cassette 1, sub-family A, member 1"
87184. 210783
/gene="Abcal"
/join(119237..119358,120364..120482,133479..133600,
135382..135758,147083..147175,149791..150031,
152367..152506,152845..152961,156570..156767,
157459..157664,158200..158376,158821..159043,
160617..160838,161790..161994,163076..163189,
164646..164817,165738..165869,167171..167313,
167492..167629,168320..168540,169790..169862,
170735..170937,172636..172684,172855..172968,
173925..174073,175605..175729,176761..176859,
180181..180370,181179..181273,182395..182427,
183398..183503,191761..191829,192306..192475,
193631..193808,196619..196734,196845..196989,
198526..198649,200922..201051,201281..201401,
202114..202176,202890..202996,204955..205096,
205427..205558,206739..206842,207361..207453,
208135..208378,209638..209778)
/gene="Abcal"
/codon_start=1
/product="ATP-binding cassette 1, sub-family A, member 1"
/protein_id="AAG39073.1"
/db_xref="GI:11611825"
/translation="MPSAGTLPWVQGIICNANPCFRYPITPGAPGVGVGNFKSIVSR
LFSDAQLLYSDRTSIKDMHVKRLQIKHPNSNLKQDLVDNETFSGFLQHLN
SLPRSTVLIQVAGQVFLQYQLOHSLASQSKLEIEIOLGDAEYSAQLGLPRK
KLDAARVLYNMILKPVYTKLNTSTPTOHLAETATVLLSLGLAQELSTSKS
NMDEEDVETDYNSTPYCNDLKNLESSPLRIKALKAPLVGLKILYTPDPATRG
VMAEVNTFQELAVFLECGMEELSPQIWFENSOEMLVETLDSRGNQDFWQK
LGLDWTADQIMFLAKNPEDVQSPGVYTRFPAFNETNOAIOITSRMECVNKLK
EPITPEVRLINKSMELDERKFWAGVITGTPDSVELPHVHYKIRMDIDNVERTNK
YKGDIFLRVMSRPMFLMTLAWISYVAIVIKVIEKEARLKETMRIMGLDGLWFS
WYFSLPLVLSAGLVILKLNLLPYSPPSVFVLSVFMVNTILOFLSTLFSR
ANLAAAGGIIYTLPLVYLVDQYVGFSGIKFASLLSVAFGFCGYFALFEEQ
GLGVQNDLFPESVEEDGNLTAVSMMLFDIFLYGWTWYIEAVPGQYGIPIRPYF
PCTKSYWFEEDIKSHPSQSQSEICMEEPHTLRLGVSTQNLVYVRDGMKVAV
DGLALNFEQGITSLFHNGAGKTTMTSLTGLPPTSGTAVILGDKDIRSESSIRQV
LGVCPQHNVLDMTVEEHIFWYARLGLSEKHVKAEMQMALDVLGPPSKLSTSQ
LSGMQRKLSVALAFYVGGKSVILDEPTAGVDPSYRGGIWEILLKYRQRTILSTHH
MDEADILGDRIATILSHKLCVGSLLFNQGLTGYYLILVKDVESLSSKNSST
VSCILKEDVSQSSDAGLGDHSDTLTIDVSAISNLIRKHVSEARLVEDIGHELY
VLPEAKAGAFVLEHEDRDLSDIGISYSETTLEIFLVAEVSVDVAETSDG
TLPARNNRFAFGDKQLHPFTDDAVDPNDSDIDPESRETLLSGMDGKGYQLKGW
KLQQQFVALLKRLILARRSKRFFAIVLPAVFCIALVFSLIWPPRGKYSLELO
PWNMYQYTFVSDNGMTNKPSPACOSDKIKMLPYCPGPGAGGLPPQRQKNTADI
LONLTGNSLDYLVKTVYQIIAKSLANKTWNEFRIGGSLGVNSQALPPSHVENDV
IKQMKLLKLTSAFRLSLGFRMAGLTDKNNKVMFNKGMHAISSFLYNNAIL
RANLQGENPSQGITAFNHLNLTQQLSEVALMTSDVLVSVICIFAMFVPASV
VFLIQRVSKAKHLQFISGRPVYIWSFNVMDCNYYVPATLVLIIFICFOQKSYV
SSTNPLVALLLLILGWSITPLMYPASVFKIPSTAYVLTVSNLFIGNGSVATFV
ELFTNNKNDINDILKSVFLIPHFCGLRGLDMVKNOAMADALFRNGRNVPSLW
DLVGRNLKAMVGVVFLITVLIOYRFFIRPRPKAKLPLNDEDEDVRRERQLID
GGQNDILEIKELTKIYRKRKPAVDICIGIPPGECFLGLVNGAGKSTFKMTGD
TPVTRDGLFNKNSILSNIEVHQNNQCPQFDATTELTGREHVEFFALLRGPKE
VGVKWAIRKGLGVKERYASNYGNGKRLSTAMALIGPPVFLDEPTTGMDPKA
RFLVNVRIAGSNPDLKPVQEFFLAPPGSVLAKRHMLOYQLPSSLSLARIFSLIS
QSKRLHIEDYSVSTLQDVFNFAKQSDDDHLKDLSLHKNQIVDVAVLTSLQD
EKWKESYV"
76801 a 58762 c 61256 g 81498 t 255 others
BASE COUNT
ORIGIN

```

```

Query Match          7.6%; Score 27; DB 10; Length 278572;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ttatctttcagtttaagaccagccac 114
|||||
Db 103522 TTTATCTTTCAGTTTAAAGCAGCCAC 103548
|||||

```

```

RESULT 38
AX239607
LOCUS AX239607 12425 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 50 from Patent WO0164874.
ACCESSION AX239607
VERSION AX239607.1 GI:15797282
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12425)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density liprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 50 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
source
location/Qualifiers
1..12425
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2563 a 3702 c 3566 g 2594 t
ORIGIN

```

```

Query Match          6.7%; Score 24; DB 6; Length 12425;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
|||||
Db 2395 GGAGGAGGAGGAGGAGGAGGAAG 2418
|||||

```

```

RESULT 39
AC006326/c
LOCUS AC006326 18427 bp DNA linear PRI 21-DEC-1999
DEFINITION Homo sapiens BAC clone GSI-234B20 from 7q31.1-q31.3, complete
sequence.
ACCESSION AC006326
VERSION AC006326.2 GI:4508133
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 18427)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 18427)
AUTHORS Maupin,R., Bauer,C. and Le,T.
TITLE The sequence of Homo sapiens BAC clone GSI-234B20
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 18427)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 18427)

```

AUTHORS TITLE JOURNAL

Waterston, R. H.
Direct Submission
Submitted (24-MAR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 18427)
Waterston, R.
Direct Submission
Submitted (01-MAY-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 18427)
Waterston, R.
Direct Submission
Submitted (10-JUL-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 18427)
Waterston, R.
Direct Submission
Submitted (21-DEC-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 24, 1999 this sequence version replaced gi:4138782.

REFERENCE AUTHORS TITLE JOURNAL

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_GS234B20

REFERENCE AUTHORS TITLE JOURNAL

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_GS234B20

REFERENCE AUTHORS TITLE JOURNAL

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_GS234B20

COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_GS234B20

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7/>, send
<mailto:edgreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone is from the first BAC library from Genome Systems, Inc.
(<http://www.genomesystems.com>).

Cell line: lymphoblastoid

Haplotypes: two

VECTOR: pBelobAC

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-866N18, 200 bp overlap;
the clone sequenced to the right is CTA-343P13, 200 bp overlap. The
actual start is unknown, the first known base of overlap is at base
position 139224 of RG114A06; actual end is at 53881 of CTA-343P13.
Location/Qualifiers
1. 18427

FEATURES source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q31.1-q31.3"
/clone="GSI-234B20"
/clone_lib="GSBAC1"

repeat_region
436. .707
/rpt_family="AT_rich"
repeat_region
708. .947
/rpt_family="Other"
misc_feature
complement(1374..3465)
/note="HNRNP A1 pseudogene"
repeat_region
1497. .1613
/rpt_family="TGG)n"
repeat_region
3523. .3646
/rpt_family="(CGGA)n"
repeat_region
3673. .3965
/rpt_family="Alu"
repeat_region
4000. .4028
/rpt_family="(GAAAA)n"
repeat_region
4029. .4162
/rpt_family="Alu"
repeat_region
4381. .4674
/rpt_family="Alu"
repeat_region
5298. .5473
/rpt_family="MER86"
repeat_region
6379. .6740
/rpt_family="MaLR"
repeat_region
7026. .7046
/rpt_family="AT_rich"
repeat_region
7246. .7544
/rpt_family="Alu"
repeat_region
7640. .7666
/rpt_family="AT_rich"
repeat_region
8042. .8132
/rpt_family="L2"
repeat_region
9524. .9818
/rpt_family="Alu"
repeat_region
10661. .10860
/rpt_family="MER1_type"
repeat_region
10955. .11122
/rpt_family="Alu"
repeat_region
11123. .11142
/rpt_family="(TAAAA)n"
repeat_region
11143. .11281
/rpt_family="Alu"
misc_feature
12018. .12543
/note="match to EST AA532765 (NID:g2278341) nf70b01.sl"
misc_feature
12068. .12557
/note="similar to Mus musculus EST AA220774 (NID:g1838594)
mv69b03.rl"
repeat_region
13097. .13385
/rpt_family="Alu"
misc_feature
13911. .14272
/note="match to EST AI033200 (NID:g3254153) ow94f08.sl"
misc_feature
14430. .14842
/note="match to EST T95217 (NID:g733841) ye44e03.rl"
misc_feature
15142. .15277
/note="match to EST AA938985 (NID:g3098898) op75f04.sl"
misc_feature
15143. .15561
/note="match to EST W93113 (NID:g1422275) zh45f06.sl"
misc_feature
15725. .16002
/note="match to EST AI202170 (NID:g3754776) qs69b06.xl"
misc_feature
15751. .16311
/note="match to EST H04382 (NID:g867315) yj21b11.sl"
misc_feature
15761. .16347
/note="match to EST W17194 (NID:g1291638) zb17a04.rl"
misc_feature
15767. .16191
/note="match to EST W93205 (NID:g1422377) zh45f06.rl"
misc_feature
15768. .16228
/note="match to EST N78828 (NID:g1241529) zb17a04.sl"
misc_feature
16018. .16276
/note="match to EST AA601910 (NID:g2436063) np02c01.sl"
misc_feature
16019. .16778
/note="match to EST AI379153 (NID:g4189006) tc66a06.xl"
misc_feature
16020. .16285
/note="similar to EST T99055 (NID:g748792) ye67a02.sl"

```

misc_feature      16021..16482
/note="match to EST AI249171 (NID:g3844568) qh68f12.xl"
misc_feature      16059..16399
/note="match to EST AA676311 (NID:g2656833) z110g04.sl"
misc_feature      16072..16571
/note="match to EST W88685 (NID:g1404396) zh71e04.rl"
misc_feature      16316..16832
/note="match to EST AI200003 (NID:g3752609) qf90c03.xl"
misc_feature      16517..16896
/note="match to EST W78848 (NID:g1389395) zh51g03.r1"
gene              complement(16664..>16891)
                  /gene="WUGSC:H_GS234B20.1"
CDS               complement(16664..>16891)
                  /gene="WNT2; match to P09544 (PID:g139750); H_GS234B20.1"
                  /codon_start=1
                  /evidence=not_experimental
                  /protein_id="AAD28351.1"
                  /db_xref="GI:4731069"
                  /translation="SLGTAGRVCMLTSGMDSCVMCCGRGYDTSHTVTRMTKCGCKFH
BASE COUNT      5121 a 3991 c 3520 g 5795 t
ORIGIN

Query Match
Best Local Similarity 6.7%; Score 24; DB 9; Length 18427;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
|||||
DB 3652 GCAGGAGGAGGAGGAGGAGGAG 3629

RESULT 40
AL583882
LOCUS
DEFINITION
Human DNA sequence from clone RP5-1098C18 on Chromosome
1p36-23-36.33, complete sequence.
ACCESSION
AL583882
VERSION
AL583882.6 GI:13990182
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90582)
Direct Submission
Submitted (03-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On May 7, 2001 this sequence version replaced gi:13897175.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:,
SWISSPROT; Tr:, TrEMBL; Wp:, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

```

Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chri>
 RP5-1098C18 is from the library RPCI-5 constructed by the group of
 Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCYPAC2
 This sequence is the entire insert of clone RP5-1098C18. The true
 right end of clone RP11-154H17 is at 1402 in this sequence.

FEATURES
 source
 Location/Qualifiers
 1..90582
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p36.23-36.33"
 /clone="RP5-1098C18"
 /clone_lib="RPCI-5"
 635..666
 /note="16 copies 2 mer tg 100% conserved"
 954..1083
 /note="13 copies 10 mer gtgtgtgtgt 66% conserved"
 982..1079
 /note="49 copies 2 mer gt 72% conserved"
 984..1079
 /note="24 copies 4 mer gtgt 72% conserved"
 1183..1286
 /note="MIR repeat: matches 41..150 of consensus"
 2755..2968
 /note="MT1B repeat: matches 1..223 of consensus"
 complement(2762..3096)
 /note="match: STS: Em:HSPE03C10"
 3462..3740
 /note="L1PB2 repeat: matches 5861..6155 of consensus"
 3890..3991
 /note="MER20 repeat: matches 118..218 of consensus"
 4018..4076
 /note="MER20 repeat: matches 1..52 of consensus"
 4909..4961
 /note="L2 repeat: matches 2693..2745 of consensus"
 5551..5610
 /note="6 copies 10 mer tacacacaca 76% conserved"
 5552..5607
 /note="14 copies 4 mer acac 75% conserved"
 5553..5610
 /note="29 copies 2 mer ca 75% conserved"
 6314..6690
 /note="L1MC3 repeat: matches 5437..5825 of consensus"
 6702..6989
 /note="AluDb repeat: matches 13..298 of consensus"
 6992..7297
 /note="AluSg repeat: matches 1..306 of consensus"
 7301..7532
 /note="AluJo repeat: matches 63..291 of consensus"
 7540..8738
 /note="L1MC3 repeat: matches 5843..7037 of consensus"
 8740..8801
 /note="Alu repeat: matches 242..298 of consensus"
 8864..9575
 /note="L1MC3 repeat: matches 7021..7738 of consensus"
 9646..10004
 /note="MT1A1 repeat: matches 1..363 of consensus"
 10015..10187
 /note="AluSp repeat: matches 132..304 of consensus"
 10338..10458
 /note="AluSp repeat: matches 1..121 of consensus"
 11856..11944
 /note="L1M4 repeat: matches 3850..3937 of consensus"
 11945..12188
 /note="MER8 repeat: matches 2..238 of consensus"
 12189..12628
 /note="L1M4 repeat: matches 3937..4387 of consensus"
 12629..12929
 /note="AluSq repeat: matches 1..301 of consensus"
 12934..13124


```

repeat_region 1055..1335 /rpt_family="AluX"
repeat_region 1055..1331 /rpt_family="AluX"
repeat_region complement(1077..1179) /rpt_family="SVA"
repeat_region 1175..1232 /rpt_family="SVA"
repeat_region 1175..1232 /rpt_family="SVA"
repeat_region 1175..1232 /rpt_family="SVA"
repeat_region 1229..1333 /rpt_family="SVA"
repeat_region 1232..1335 /rpt_family="SVA"
repeat_region 1272..1346 /rpt_family="SVA"
repeat_region /note="SST_REPEAT"
repeat_region 1272..1346 /note="SST_REPEAT"
repeat_region /note="XNUN_REPEAT"
repeat_region complement(1332..1628) /rpt_family="AluX"
repeat_region complement(1339..1627) /rpt_family="AluX"
repeat_region complement(1339..1628) /rpt_family="AluX"
repeat_region 1347..1628 /rpt_family="AluX"
repeat_region /rpt_type="inverted"
repeat_region complement(1366..1423) /rpt_family="SVA"
repeat_region complement(1410..1474) /rpt_family="SVA"
repeat_region complement(1412..1488) /rpt_family="SVA"
repeat_region complement(1431..1515) /rpt_family="SVA"
repeat_region complement(1502..1605) /rpt_family="SVA"
repeat_region 1506..1605 /rpt_family="SVA"
repeat_region complement(1535..1628) /rpt_family="SVA"
repeat_region complement(1566..1628) /rpt_family="SVA"
repeat_region 1624..1730 /rpt_family="SVA"
repeat_region /rpt_family="MIR"
repeat_region complement(1624..1737) /rpt_family="MIR"
repeat_region 1629..1737 /rpt_family="MIR"
repeat_region 1796..1841 /note="XNUN_REPEAT"
repeat_region 1799..1841 /note="SST_REPEAT"
exon complement(1800..1985) /note="GRAIL"
repeat_region /evidence="not_experimental"
repeat_region 1966..2073 /note="SST_REPEAT"
repeat_region 2756..2847 /rpt_family="AluX"
repeat_region 2756..2824 /rpt_family="SVA"
repeat_region 2758..3051 /rpt_family="AluX"
repeat_region 2758..2869 /rpt_family="AluX"
repeat_region 2777..2872 /rpt_family="SVA"
repeat_region complement(2777..2868) /rpt_family="SVA"
exon complement(2822..3037) /note="FXHB"
repeat_region /evidence="not_experimental"
repeat_region 2863..3047 /rpt_family="AluX"
repeat_region 2867..3048 /rpt_family="AluX"
repeat_region 2883..2954 /rpt_family="SVA"
repeat_region 2894..2987 /rpt_family="SVA"
repeat_region 2913..2976 /rpt_family="SVA"
repeat_region 2962..3017 /rpt_family="SVA"
repeat_region 3041..3105 /note="SST_REPEAT"
repeat_region 3057..3208 /rpt_family="Llmc3"
repeat_region complement(3070..3271) /rpt_family="Llmc3"
repeat_region 3137..3205 /rpt_family="Llmc4"
repeat_region 3138..3205 /rpt_family="Llmc2"
repeat_region complement(3142..3207) /rpt_family="Llmc3"
repeat_region complement(3371..3421) /rpt_family="MIR"
repeat_region 3372..3545 /rpt_family="MIR"
repeat_region complement(3405..3461) /rpt_family="MIR"
repeat_region complement(3721..3802) /rpt_family="MIR"
repeat_region 3762..3900 /rpt_family="MIR"
repeat_region 4007..4055 /note="SST_REPEAT"
repeat_region 4108..4318 /rpt_family="MSTB"
repeat_region 4108..4246 /rpt_family="MSTA"
repeat_region complement(4110..4245) /rpt_family="MSTA"
repeat_region complement(4154..4218) /rpt_family="the1b"
exon 4211..4246 /note="XPOUND"
repeat_region /evidence="not_experimental"
repeat_region 4290..4414 /rpt_family="AluX"
repeat_region 4319..4619 /rpt_family="AluSg"
repeat_region 4322..4449 /rpt_family="SVA"
repeat_region 4322..4491 /rpt_family="SVA"
repeat_region 4332..4484 /rpt_family="AluX"
repeat_region complement(4342..4443) /rpt_family="SVA"
repeat_region 4345..4706 /rpt_type="inverted"
repeat_region 4347..4706 /rpt_type="inverted"
repeat_region 4349..4706 /rpt_type="inverted"
repeat_region 4351..4417 /rpt_type="inverted"
repeat_region 4465..4606 /rpt_family="AluX"
repeat_region 4524..4579 /rpt_family="SVA"
repeat_region 4709..5070 /rpt_type="inverted"
repeat_region 4709..5066 /rpt_type="inverted"

```


repeat_region /rpt_type=inverted 4709..5068
repeat_region /rpt_type=inverted 4775..5004
repeat_region /note="XNON_REPEAT" 4775..4866
repeat_region /rpt_type=tandem complement(4775..4866)
/note="(TGGA)n"
/rpt_family="simple_repeat"

Query Match 6.7% Score 24; DB 2; Length 123339;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggagggaggaggaggaggagg 349
|||||
Db 82626 GGAGGAGGGAGGGAGGGAGGAGGAGG 82603

RESULT 42
AC036144/c 127270 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens chromosome 8, clone RP11-29316, complete sequence.
DEFINITION AC036144
ACCESSION AC036144.7 GI:18104887
VERSION HTG.
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 127270)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-29316
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 127270)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Gallagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Rile,R., Rogov,P.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rose,C., Rogov,P.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (07-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 127270)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (07-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 127270)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (18-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 127270)
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (18-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 127270)
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (18-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 127270)
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L8779
 Center clone name: 293_L_6

FEATURES
 source T7 end overlaps AC009941 [WTCGR Project L1987] by 103216 bp.

```

Location/Qualifiers
1. .127270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-293L6"
/clone_lib="RPC1-11 Human Male BAC"
complement(4. 3746)
/rpt_family="L1"
946. .953
/note="<30 qual SNGL region"
1029. .1033
/note="<30 qual SNGL region"
1056. .1060
/note="<30 qual SNGL region"
1110. .1126
/note="<30 qual SNGL region"
1135. .1141
/note="<30 qual SNGL region"
1145. .1152
/note="<30 qual SNGL region"
1169. .1177
/note="<30 qual SNGL region"
1190. .1195
/note="<30 qual SNGL region"
1219. .1224
/note="<30 qual SNGL region"
1230. .1234
/note="<30 qual SNGL region"
2097. .2102
/note="<30 qual SNGL region"
complement(3768. .3859)
/rpt_family="L1"
complement(3860. .4173)
/rpt_family="AluY"
complement(4179. .6609)
/rpt_family="L1PA7"
6610. .6647
/rpt_family="(TTG)n"
complement(6648. .8549)
/rpt_family="L1PA7"
complement(8549. .9721)
/rpt_family="L1"
complement(9736. .10077)
/rpt_family="L1"
10180. .10236
/rpt_family="AT_rich"
complement(10474. .10599)
/rpt_family="MIR"
10698. .10800
/rpt_family="MIRNA"
11389. .11416
/rpt_family="AT_rich"
complement(12357. .12484)
/rpt_family="MIR3"
complement(12621. .12742)
/rpt_family="L2"
complement(12759. .13088)
/rpt_family="MER7A"
complement(13108. .13281)
/rpt_family="MER5B"
complement(13285. .13379)
/rpt_family="MER5B"
14034. .14057
/rpt_family="AT_rich"

```

```

repeat_region complement(14859. .14987)
repeat_region /rpt_family="L2"
repeat_region 15347. .15525

Query Match 6.7%; Score 24; DB 9; Length 127270;
Best Local Similarity 100.0%; Pred. NO. 0.024;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
DB 18870 GGAGGAGGAGGAGGAGGAGGAGG 18847

RESULT 43
ALI58043/c
LOCUS
DEFINITION
ALI58043
Human DNA sequence from clone RP11-5N23 on chromosome 10p14-15.3
Contains ESTs, GSSs, STSS and a CpG island. Contains the 5' part of
the PRKQ gene for protein kinase C theta and a novel gene,
complete sequence.
ALI58043
ALI58043.14 GI:9863558
HTG; CpG island; PRKQ; protein kinase.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 132741)
AUTHORS
Chapman,J.
TITLE
Direct Submission
JOURNAL
Submitted (09-OCT-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9856693.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormep/ This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-5N23 is from the library RPC1-11.1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-5N23 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-5N23 is at 132741 in this
sequence. The true left end of clone RP11-55418 is at 102011 in
this sequence. The true right end of clone RP11-563J2 is at 100 in
this sequence.
FEATURES
source Location/Qualifiers
1. .132741
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="p14-15.3"
/clone="RP11-5N23"

```



```

QY 326 ggaggaggaggaggaggaggagg 349
|||||
Db 1980 GGAGGAGGAGGAGGAGGAGGAGG 1957

RESULT 44
AL163512
LOCUS
DEFINITION
  Mouse DNA sequence from clone RP21-468E12 on chromosome X, complete
  sequence.
ACCESSION
  AL163512
VERSION
  AL163512.35 GI:17017692
KEYWORDS
  HTG.
SOURCE
  house mouse.
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 134334)
REFERENCE
  Wall,M.
  Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk
  On Nov 20, 2001 this sequence version replaced gi:15808147.
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone.
  only a small overlap as described above.
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
  one plasmid subclone or more than one M13 subclone; and the
  assembly was confirmed by restriction digest. The following
  abbreviations are used to associate primary accession numbers given
  in the feature table with their source databases: Em, EMBL; SW,
  SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
  database can be found at
  http://www.sanger.ac.uk/projects/C_elegans/wormpep RP21-468E12 is
  from the RP21-21 Mouse PAC Library
  constructed by the group of Pieter de Jong.
  For further details see http://www.chori.org/bacpac/home.htm
  VECTOR: pPAC4
  This sequence is the entire insert of clone RP21-468E12.
  Location/Qualifiers
    1..134334
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /chromosome="X"
      /clone="RP21-468E12"
      /clone_lib="RPCI-21"
      15023
misc_feature
  /note="Tandem repeat. Forced join, Gap size estimated to
  be approximately 280bp by restriction digest data."
  114512
misc_feature
  /note="Tandem repeat. Forced join, Gap size estimated to
  be approximately 90bp by restriction digest data."
  complement(134059..134062)
misc_feature
  /note="Single clone region. Assembly confirmed by
  restriction digest data."
  complement(134064..134068)
misc_feature
  /note="Single clone region. Assembly confirmed by
  restriction digest data."
  complement(134070..134075)
misc_feature
  /note="Single clone region. Assembly confirmed by
  restriction digest data."
  complement(134077..134087)
misc_feature
  /note="Single clone region. Assembly confirmed by
  restriction digest data."

QY 326 ggaggaggaggaggaggaggagg 349
|||||
Db 114497 GGAGGAGGAGGAGGAGGAGGAGG 114520

RESULT 45
AL353893
LOCUS
DEFINITION
  Homo sapiens chromosome 1 clone RP5-1108E5, *** SEQUENCING IN
  PROGRESS ***, 6 unordered pieces.
ACCESSION
  AL353893
VERSION
  AL353893.2 GI:9213229
KEYWORDS
  HTG; HTGS_PHASE1; HTGS-CANCELLED.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 137072)
REFERENCE
  McIay,K.
  Direct Submission
  Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
  requests: clonerequest@sanger.ac.uk
  On Jul 15, 2000 this sequence version replaced gi:8052073.
  ----- Genome Center
  Center: Sanger Centre
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: humquery@sanger.ac.uk
  ----- Project Information
  Center project name: dj1108E5
  ----- Summary Statistics
  Assembly program: XGAP4; version 4.5
  Sequencing vector: plasmid; L08752; 100% of reads
  Chemistry: dye-terminator Big Dye; 100% of reads
  Consensus quality: 134274 bases at least Q40
  Consensus quality: 135563 bases at least Q30
  Consensus quality: 136185 bases at least Q20
  Insert size: 136572; sum-of-contigs
  Quality coverage: 4.52x in Q20 bases; agarose-fp
  coverage: 4.50x in Q20 bases; agarose-fp
  -----
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 6 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  *
  * 1 4285: contig of 4285 bp in length
  * 4286 4385: gap of 100 bp
  * 4386 67924: contig of 63539 bp in length
  * 67925 68024: gap of 100 bp
  * 68025 71010: contig of 2986 bp in length
  * 71011 71110: gap of 100 bp
  * 71111 89837: contig of 18727 bp in length
  * 89838 89937: gap of 100 bp
  * 89938 113010: contig of 23073 bp in length
  * 113011 113110: gap of 100 bp
  * 11311 137072: contig of 23962 bp in length.
  Location/Qualifiers
    1..137072
      /organism="Homo sapiens"
FEATURES
  source
    1..137072

```

```

/db_xref="taxon:9606"
/chromosome="1"
/clone="RP5-1108E5"
/clone_lib="RPC1-5"
1..4285
/note="assembly_fragment:00373
fragment_chain:1"
4386..67924
/note="assembly_fragment:00570
fragment_chain:1"
68025..71010
/note="assembly_fragment:00662
fragment_chain:2"
71111..89837
/note="assembly_fragment:00832
fragment_chain:2"
89938..113010
/note="assembly_fragment:00769"
113111..137072
/note="assembly_fragment:01179"
BASE COUNT 38932 a 29307 c 29340 g 38993 t 500 others
ORIGIN

```

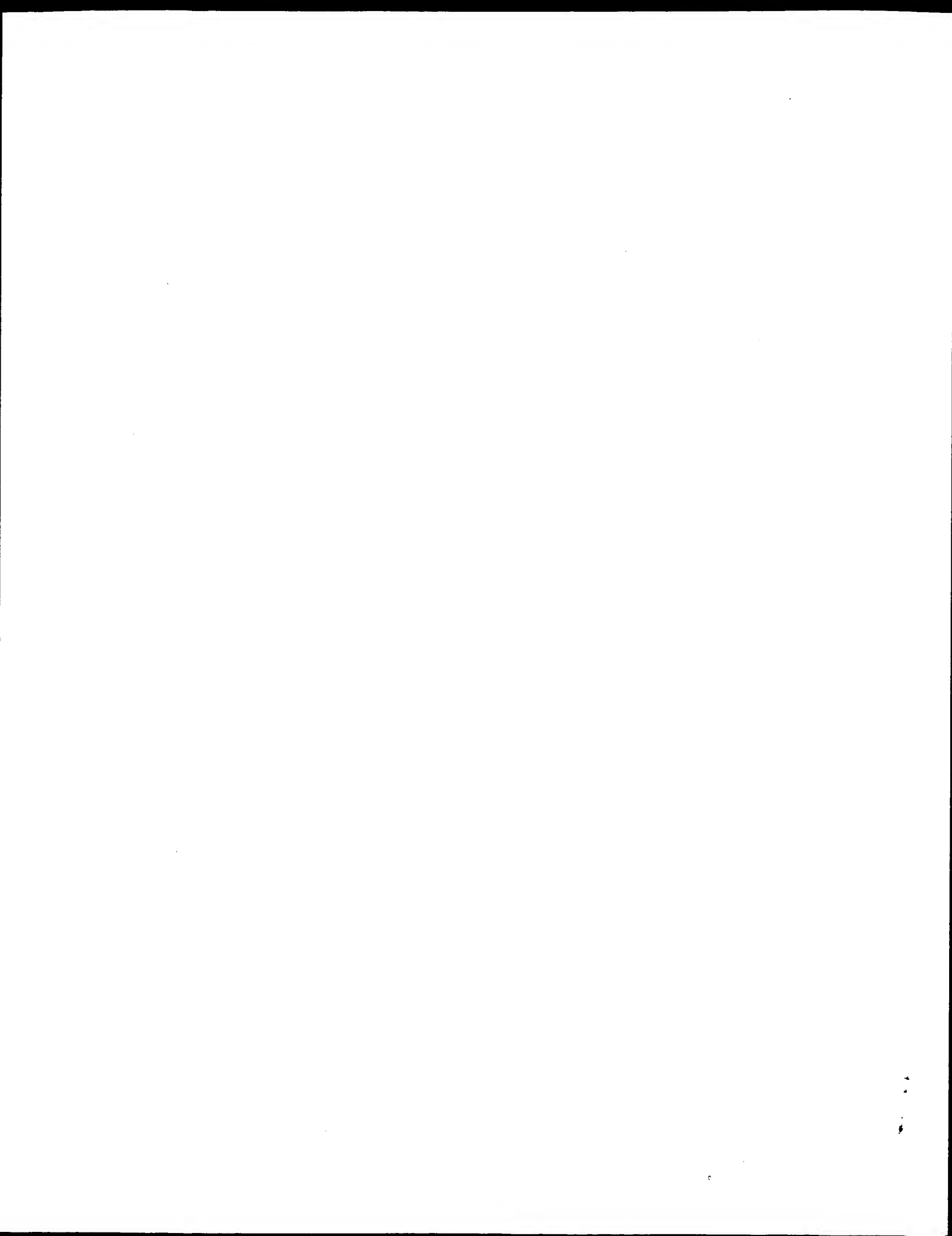
```

Query Match      6.7%; Score 24; DB 2; Length 137072;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
Db 22555 GGAGGAGGGAGGGAGGGAAG 22578

```

Search completed: September 20, 2002, 06:30:00
Job time: 18394 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:06:55 ; Search time 521.76 Seconds
(without alignments)
1174.752 Million cell updates/sec

Title: US-09-846-456-2
Perfect score: 357
Sequence: 1 tggaggtctcagctgagagg.....gagggaaggaagctgtgttg 357

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	159	44.5	10545	21 AAC69132	Human ABC1 gene ex
2	159	44.5	183999	22 AAF92831	Human ABC1 genomic
3	77	21.6	10442	22 AAF24680	Nucleotide sequenc
4	77	21.6	10442	22 AAF24702	Nucleotide sequenc
5	77	21.6	10474	22 AAF24685	Nucleotide sequenc
6	77	21.6	10474	22 AAF24686	Nucleotide sequenc
7	77	21.6	10474	22 AAF24707	Nucleotide sequenc
8	77	21.6	10474	22 AAF24708	Nucleotide sequenc
9	60	16.8	446	22 AAS04035	Partial human ABC1

10	60	16.8	7086	22 ABA03200	Human ABCA1 homolo
11	60	16.8	7086	22 AAK52667	Human polynucleoti
12	60	16.8	7260	22 AAD21326	Human ATP binding
13	60	16.8	7260	22 AAI70315	Human ATP binding
14	60	16.8	7281	22 AAK51683	Human polynucleoti
15	60	16.8	7857	21 AAC69388	Human ABC1 choles
16	60	16.8	7860	22 AAF83826	Human ABC1 nucleot
17	60	16.8	7860	22 AAF92835	Human ABC1 cDNA.
18	60	16.8	7861	21 AAC69387	Human ABC1 choles
19	60	16.8	7864	21 AAC69120	Human ABC1 choles
20	60	16.8	7864	21 AAC69385	Human ABC1 choles
21	60	16.8	7864	21 AAC69386	Human ABC1 choles
22	60	16.8	7864	21 AAC69389	Human ABC1 choles
23	60	16.8	9741	22 AAS06120	Human ABC1 DNA seq
24	60	16.8	9854	22 AAS06121	Human ABC1 DNA seq
25	51	14.3	736	22 AAH07432	Human cDNA clone (
26	51	14.3	1556	22 AAH18606	Human cDNA sequenc
27	35	9.8	37	22 AAF93084	ABC1 polymorphism
28	35	9.8	38	22 AAF93082	Human low density
29	24	6.7	12425	22 AAF26495	Human BMP-4 5' ups
30	22	6.2	30	21 AAZ39175	DNA sequence from
31	22	6.2	1077	22 AAC83229	Human BMP-4 promot
32	22	6.2	1456	19 AAV32783	BMP4. Rattus ratt
33	22	6.2	1751	13 AAQ32853	Bone morphogenetic
34	22	6.2	1751	19 AAV01680	Sequence encoding
35	22	6.2	1788	14 AAQ53144	Prepro human CBMP2
36	22	6.2	1788	15 AAQ72710	Human CBMP2(b) CDN
37	22	6.2	1788	17 AAT02601	Human osteogenic p
38	22	6.2	1788	19 AAV15207	Human prepro CBMP2
39	22	6.2	1788	20 AAZ27580	Human osteogenic p
40	22	6.2	1832	23 AAX00232	DNA encoding novel
41	22	6.2	1832	23 AAS79379	DNA encoding novel
42	22	6.2	1944	18 AAT78942	Human bone morphog
43	22	6.2	1954	9 AAN80634	Human BMP-2B in la
44	22	6.2	1954	12 AAQ14037	Human BMP-2B sequ
45	22	6.2	1954	13 AAQ31870	

ALIGNMENTS

RESULT 1	
AAC69132	
ID AAC69132 standard; DNA; 10545 BP.	
XX	
AC AAC69132;	
XX	
DT 29-JAN-2001 (first entry)	
XX	
DE Human ABC1 gene exon 1 (promoter).	
XX	
KW Human ABC1 cholesterol transporter; chromosome 9q31; promoter;	
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;	
KW Tangier disease; TB; familial HDL deficiency; FHA; polymorphism;	
KW cardiovascular disease; coronary artery disease; coronary restenosis;	
KW cerebrovascular disease; peripheral vascular disease;	
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;	
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;	
KW prognosis; prophylaxis; drug screening; transgenic animal; ss.	
XX	
OS Homo sapiens.	
XX	
PN WO200055318-A2.	
XX	
PD 21-SEP-2000.	
XX	
PF 15-MAR-2000; 2000WO-IB00532.	
XX	
PR 15-MAR-1999; 99US-0124702.	
PR 08-JUN-1999; 99US-0138048.	
PR 17-JUN-1999; 99US-0139600.	
PR 01-SEP-1999; 99US-0151977.	
XX	

QY 191 acatggcatgttgccctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 250
 Db 289 acatggcttggctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 348
 QY 251 gacaaaca 258
 Db 349 gacaaaca 356
 RESULT 5
 ID AAF24685
 XX AAF24685 standard; DNA; 10474 BP.
 AC AAF24685;
 XX
 DT 20-APR-2001 (first entry)
 XX
 Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 CDS Location/Qualifiers
 FT 323..7108
 FT /*tag= a
 FT /product= "defective ABC1 polypeptide"
 XX
 PN W0200078972-A2.
 XX
 PD 28-DEC-2000.
 XX
 PF 16-JUN-2000; 2000WO-US16765.
 XX
 PR 18-JUN-1999; 99US-0140264.
 PR 14-SEP-1999; 99US-0153872.
 PR 19-NOV-1999; 99US-0166573.
 XX
 PA (CVTH-) CV THERAPEUTICS INC.
 XX
 PI Lawn RM, Wade D, Garvin M;
 XX
 DR WPI; 2001-137812/14.
 XX
 PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
 PT useful for the development of agents for the treatment of heart disease
 PT and other disorders associated with hypercholesterolemia and
 PT atherosclerosis -
 XX
 PS Disclosure; Page 148-154; 215pp; English.
 XX
 CC The present sequence encodes a human adenosine triphosphate (ATP)
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
 CC ATP hydrolysis to transport a wide variety of substrates across the
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
 CC Tangier disease, a genetic disorder characterised by abnormal
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
 CC 9q22-q31. The ABC1 genes and proteins are useful for developing
 CC pharmaceutical agents for the treatment of heart disease and other
 CC disorders associated with hypercholesterolemia and atherosclerosis. The
 CC genes are useful for developing screening assays to screen for compounds
 CC that regulate the expression of genes associated with cholesterol
 CC transport. The genes and proteins are also useful for are also useful
 CC as diagnostic indicators of cardiovascular disease and other disorders
 CC associated with hypercholesterolemia.
 XX
 SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match 21.6%; Score 77; DB 22; Length 10474;
 Best Local Similarity 99.2%; Pred. No. 4.8e-28;
 Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 131 agctctggccgctgcctccagggctccgagccacacacgctggcgctgctgagggga 190
 Db 261 agctctggccgctgcctccagggctccgagccacacacgctggcgctgctgagggga 320
 QY 191 acatggcatgttgccctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 250
 Db 321 acatggcttggctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 380
 QY 251 gacaaaca 258
 Db 381 gacaaaca 388
 RESULT 6
 ID AAF24686
 XX AAF24686 standard; DNA; 10474 BP.
 AC AAF24686;
 XX
 DT 20-APR-2001 (first entry)
 XX
 Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 CDS Location/Qualifiers
 FT 323..7108
 FT /*tag= a
 FT /product= "defective ABC1 polypeptide"
 XX
 PN W0200078972-A2.
 XX
 PD 28-DEC-2000.
 XX
 PF 16-JUN-2000; 2000WO-US16765.
 XX
 PR 18-JUN-1999; 99US-0140264.
 PR 14-SEP-1999; 99US-0153872.
 PR 19-NOV-1999; 99US-0166573.
 XX
 PA (CVTH-) CV THERAPEUTICS INC.
 XX
 PI Lawn RM, Wade D, Garvin M;
 XX
 DR WPI; 2001-137812/14.
 XX
 PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
 PT useful for the development of agents for the treatment of heart disease
 PT and other disorders associated with hypercholesterolemia and
 PT atherosclerosis -
 XX
 PS Disclosure; Page 170-176; 215pp; English.
 XX
 CC The present sequence encodes a human adenosine triphosphate (ATP)
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
 CC ATP hydrolysis to transport a wide variety of substrates across the
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
 CC Tangier disease, a genetic disorder characterised by abnormal
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
 CC 9q22-q31. The ABC1 genes and proteins are useful for developing
 CC pharmaceutical agents for the treatment of heart disease and other

CC disorders associated with hypercholesterolemia and atherosclerosis. The
CC genes are useful for developing screening assays to screen for compounds
CC that regulate the expression of genes associated with cholesterol
CC transport. The genes and proteins are also useful for are also useful
CC as diagnostic indicators of cardiovascular disease and other disorders
CC associated with hypercholesterolemia.
XX
SQ Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

Query Match 21.6%; Score 77; DB 22; Length 10474;
Best Local Similarity 99.2%; Pred. No. 4.8e-28;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 131 agctctggcgcgtgctccagggctcccgagccacacgctggcgctgctggctgagga 190
DB 261 agctctggcgcgtgctccagggctcccgagccacacgctggcgctgctggctgagga 320

QY 191 acatggcatgttgctcagctgaggtgtgctggaagaacccctcactttcagaagaa 250
DB 321 acatggctgttgctcagctgaggtgtgctggaagaacccctcactttcagaagaa 380

QY 251 gacaaaca 258
DB 381 gacaaaca 388

RESULT 7
AAF24707
ID AAF24707 standard; DNA; 10474 BP.
XX
AC AAF24707;
XX
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 323..7108
FT /*tag= a
FT /product= "defective ABC1 polypeptide"
XX
XX WO200078971-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16591.
XX
XX 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
XX 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
PA (UNIW) UNIV WASHINGTON.
XX
XX Lawn RM, Wade D, Oram JF, Garvin M;
PI WPI; 2001-137811/14.
XX
XX P-PSDB; AAB31366.
XX
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT polynucleotides and polypeptides, useful for treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
XX Claim 27; Page 144-150; 211pp; English.

XX The present sequence encodes a human adenosine triphosphate (ATP)
CC binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC ATP hydrolysis to transport a wide variety of substrates across the
CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC Tangier disease, a genetic disorder characterised by abnormal
CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC 9q22-9q31. The ABC1 genes and proteins are useful for developing
CC pharmaceutical agents for the treatment of heart disease and other
CC disorders associated with hypercholesterolemia and atherosclerosis. The
CC genes are useful for developing screening assays to screen for compounds
CC that regulate the expression of genes associated with cholesterol
CC transport. The genes and proteins are also useful for are also useful
CC as diagnostic indicators of cardiovascular disease and other disorders
CC associated with hypercholesterolemia.
XX
SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match 21.6%; Score 77; DB 22; Length 10474;
Best Local Similarity 99.2%; Pred. No. 4.8e-28;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 131 agctctggcgcgtgctccagggctcccgagccacacgctggcgctgctggctgagga 190
DB 261 agctctggcgcgtgctccagggctcccgagccacacgctggcgctgctggctgagga 320

QY 191 acatggcatgttgctcagctgaggtgtgctggaagaacccctcactttcagaagaa 250
DB 321 acatggctgttgctcagctgaggtgtgctggaagaacccctcactttcagaagaa 380

QY 251 gacaaaca 258
DB 381 gacaaaca 388

RESULT 8
AAF24708
ID AAF24708 standard; DNA; 10474 BP.
XX
AC AAF24708;
XX
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 323..7108
FT /*tag= a
FT /product= "defective ABC1 polypeptide"
XX
XX WO200078971-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16591.
XX
XX 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
XX 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
PA (UNIW) UNIV WASHINGTON.
XX
XX

PD 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US03800.
 PF 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX (HYSE-) HYSEQ INC.
 PA Tang Yt, Liu C, Drmanac RT;
 PI WPI; 2001-457740/49.
 DR P-PSDB: ABB11956.
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX Claim 1; Page 833-835; 1963pp; English.
 PS Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, and hence
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoietic regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis; cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.
 XX
 SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7086;
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258
 Db 310 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 369

RESULT 11
 AAK52667
 ID AAK52667 standard; cDNA; 7086 BP.

XX AAK52667;
 AC 06-NOV-2001 (first entry)
 XX Human polynucleotide SEQ ID NO 2196.
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX Homo sapiens.
 OS WO200157190-A2.
 XX 09-AUG-2001.
 PD 05-FEB-2001; 2001WO-US04098.
 PF 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PA Tang Yt, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 DR WPI; 2001-476283/51.
 DR P-PSDB: AAM79534.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 1; Page 4558-4560; 6221pp; English.
 PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7086;
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258
 Db 310 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 369

RESULT 12
 AAD21326
 ID AAD21326 standard; DNA; 7260 BP.

Query Match 16.8%; Score 60; DB 22; Length 7260;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggtgctgtgtggaagaacctcacttctcagaagaacaaaca 258
|||||
Db 327 tttggcctcagctgaggtgctgtgtggaagaacctcacttctcagaagaacaaaca 386
|||||

RESULT 14

AAK51683
ID AAK51683 standard; cDNA; 7281 BP.

XX AC AAK51683;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 228.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX P-PSDB; AAK78550.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -

XX PS Claim 1; Page 1086-1096; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAK78323-AAK80302) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.

XX CC Note: records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAK80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.

XX SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7281;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggtgctgtgtggaagaacctcacttctcagaagaacaaaca 258
|||||
Db 348 tttggcctcagctgaggtgctgtgtggaagaacctcacttctcagaagaacaaaca 407
|||||

RESULT 15

AAK69388
ID AAK69388 standard; cDNA; 7857 BP.

XX AC AAK69388;

XX DT 29-JAN-2001 (first entry)

XX DE Human ABC1 cholesterol transporter FHA-3 mutant cDNA (delta 5752-5757).

XX KW Human ABC1 cholesterol transporter; chromosome 9q31;

XX KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;

XX KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;

XX KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;

XX KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.

XX OS Homo sapiens.

XX PN WO2000055318-A2.

XX PD 21-SEP-2000.

XX PR 15-MAR-2000; 2000WO-IB00532.

XX PR 15-MAR-1999; 99US-0124702.

XX PR 08-JUN-1999; 99US-0138048.

XX PR 17-JUN-1999; 99US-0139600.

XX PR 01-SEP-1999; 99US-0151977.

XX PA (UYBR-) UNIV BRITISH COLUMBIA.

XX PA (XENO-) XENON BIORESEARCH INC.

XX PI Hayden MR, Wilson AR, Pimstone SN;

XX WPI; 2000-587528/55.

XX P-PSDB; AAB38107.

XX PT New ABC1 polypeptide is useful for treating diseases associated with
XX ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
XX disease and cancer -

XX PS Examples; Page -; 229pp; English.

XX CC The invention relates to the human ABC1 cholesterol transporter protein
XX (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
XX a member of the ATP-binding cassette (ABC transporter) superfamily of
XX proteins, and plays a crucial role in cholesterol transport, particularly
XX intracellular cholesterol trafficking in monocytes and fibroblasts, being
XX involved in cholesterol efflux from the cell. The gene encoding ABC1 is
XX located on chromosome 9q31, and mutations in this gene are associated
XX with two genetic HDL (high density lipoprotein) deficiency disorders,
XX Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
XX are distinguishable in that TD is an autosomal recessive disorder, while
XX FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
XX cholesterol") in the blood correlate with a high risk of cardiovascular
XX disease, particularly coronary artery disease, but also cerebrovascular
XX disease, coronary restenosis, and peripheral vascular disease.
XX Conversely, a high level of HDL has protective effects against
XX cardiovascular disease. The invention provides genetic constructs and

transgenic cells and non-human animals comprising human ABC1 nucleic acids, and methods of gene therapy for the treatment or prevention of cardiovascular disease comprising the administration of an expression vector encoding ABC1 or an active fragment thereof. The invention also encompasses compounds which mimic ABC1 activity, compounds which stimulate ABC1 expression and methods of screening for such compounds. It further relates to methods for determining whether a patient has an increased risk for cardiovascular disease due to polymorphisms in the ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or prevent cardiovascular disease, especially coronary artery disease, cerebrovascular disease, coronary restenosis or peripheral vascular disease. They may also be used in the treatment of diseases associated with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer. The invention specifically excludes proteins with the exact amino acid sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic acid with the exact sequence as GenBank Accession No: AJ012376.1. The present sequence represents cDNA encoding a mutant human ABC1 cholesterol transporter associated with an altered cholesterol level and therefore an altered risk of cardiovascular disease.

Note: The present sequence is not shown in the specification, but is derived from the native human ABC1 cDNA shown on pages 157-160.

Sequence 7857 BP; 2011 A; 1860 C; 2008 G; 1977 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7857;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 258
|||||
Db 81 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 140

RESULT 16
AAF83826
ID AAF83826 standard; DNA; 7860 BP.
XX
AC AAF83826;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human ABC1 nucleotide sequence.
XX
KW ABC1; antilipemic; cholesterol; inhibitor; low density lipoprotein;
XX LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 75..3341
FT /*tag= a
FT /product= "partial ABC1 protein"
FT /note= "the coding sequence continues beyond nucleotide
3341, possibly till position 6860 as identified
by translating the present sequence; part of the
corresponding protein is missing and nucleotide
3341 corresponds to the last amino acid residue
(position 1089) as indicated in the
specification"

WO2001132184-A2.
XX
PN Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX
PD WPI; 2001-244356/25.
XX
XX
XX
PF 01-NOV-2000; 2000WO-US30109.
XX
XX
PR 01-NOV-1999; 99US-0162803.
PR 30-JUN-2000; 2000US-0215564.
XX
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX

PI Attie AD, Cook M, Gray-Keller MP, Hayden MR, Pimstone S;
PI Brooks-Wilson A;
XX
XX WPI; 2001-335779/35.
DR P-PSDB; AAB62691.
DR
XX
PT New method for inhibiting cholesterol uptake in the gut comprises
PT administration of an inhibitor of an ABC1 protein -
XX
XX Disclosure; Page 34-36; 41pp; English.
XX
CC The invention relates to a new method for inhibiting cholesterol uptake
CC in the gut that comprises administration of an inhibitor of an ABC1
CC protein. The method is useful for: lowering levels of LDL (low density
CC lipoprotein) cholesterol by reducing the activity of ABC1 protein in the
CC intestinal cells and the abundance of the ABC1 protein in the individual
CC by inhibiting the activity of the protein; identifying drugs that can
CC lower serum cholesterol and LDL levels comprising assaying the drug to
CC test if it can bind to an ABC1 protein; testing LDL cholesterol lowering
CC agents; and for modulation of ABC1 biological activity. The present
CC sequence represents a human ABC1 nucleotide sequence.

Sequence 7860 BP; 2013 A; 1861 C; 2009 G; 1977 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7860;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 258
|||||
Db 81 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 140

RESULT 17
AAF92835
ID AAF92835 standard; DNA; 7860 BP.
XX
AC AAF92835;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human ABC1 cDNA.
XX
KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ss.
XX
OS Homo sapiens.
XX
PN WO200115676-A2.
XX
PD 08-MAR-2001.
XX
XX
PF 01-SEP-2000; 2000WO-IB01492.
XX
XX
PR 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
XX
PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX
DR WPI; 2001-244356/25.
XX
XX
PT Treating a lower than normal high density lipoprotein-cholesterol
PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
XX Disclosure; Fig 2; 317pp; English.
XX
XX The present invention relates to a method for treating a patient

diagnosed as having a lower than normal high density lipoprotein-cholesterol (HDL-C) level, a higher than normal triglyceride level, or a cardiovascular disease, involving administering a compound that modulates LXR- or RXR-mediated transcriptional activity or ABC1 expression or activity.

The LXR gene product may be used in an assay to identify compounds useful for the treatment of a disease or condition selected a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

Sequence 7860 BP; 2014 A; 1860 C; 2008 G; 1978 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7860; Best Local Similarity 100.0%; Pred. No. 1.1e-19; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

QY 199 ttttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 258
|||||
Db 81 ttttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 18
AAC69387
ID AAC69387 standard; cDNA; 7861 BP.
AC AAC69387;
XX
XX
XX 29-JAN-2001 (first entry)
XX Human ABC1 cholesterol transporter FHA-1 mutant cDNA (delta 2151-2153).
XX
XX Human ABC1 cholesterol transporter; chromosome 9q31;
XX ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
XX Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
XX cardiovascular disease; coronary artery disease; coronary restenosis;
XX cerebrovascular disease; peripheral vascular disease;
XX Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
XX X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
XX prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
XX Homo sapiens.
XX OS
XX WO200055318-A2.
XX PN
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-IB00532.
XX
XX 15-JUN-1999; 99US-0124702.
XX PR 08-JUN-1999; 99US-0138048.
XX PR 17-JUN-1999; 99US-0139600.
XX PR 01-SEP-1999; 99US-0151977.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX (XENO-) XENON BIORESEARCH INC.
XX PA
XX Hayden MR, Wilson AR, Pimstone SN;
XX PI
XX WPI; 2000-587528/55.
XX DR P-PSDB; AAB38106.
XX
XX New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
PT
PT
XX Examples; Page -; 229pp; English.
XX PS
XX The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being

involved in cholesterol efflux from the cell. The gene encoding ABC1 is located on chromosome 9q31, and mutations in this gene are associated with two genetic HDL (high density lipoprotein) deficiency disorders. Tangier disease (TD) and familial HDL deficiency (FHA). These diseases are distinguishable in that TD is an autosomal recessive disorder, while FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good cholesterol") in the blood correlate with a high risk of cardiovascular disease, particularly coronary artery disease, but also cerebrovascular disease, coronary restenosis, and peripheral vascular disease. Conversely, a high level of HDL has protective effects against cardiovascular disease. The invention provides genetic constructs and transgenic cells and non-human animals comprising human ABC1 nucleic acids, and methods of gene therapy for the treatment or prevention of cardiovascular disease comprising the administration of an expression vector encoding ABC1 or an active fragment thereof. The invention also encompasses compounds which mimic ABC1 activity, compounds which stimulate ABC1 expression and methods of screening for such compounds. It further relates to methods for determining whether a patient has an increased risk for cardiovascular disease due to polymorphisms in the ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or prevent cardiovascular disease, especially coronary artery disease, cerebrovascular disease, coronary restenosis or peripheral vascular disease. They may also be used in the treatment of diseases associated with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer. The invention specifically excludes proteins with the exact amino acid sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic acid with the exact sequence as GenBank Accession No: AJ012376.1. The present sequence represents cDNA encoding a mutant human ABC1 cholesterol transporter associated with an altered cholesterol level and therefore an altered risk of cardiovascular disease.

Note: The present sequence is not shown in the specification, but is derived from the native human ABC1 cDNA shown on pages 157-160.

Sequence 7861 BP; 2014 A; 1859 C; 2011 G; 1976 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7861; Best Local Similarity 100.0%; Pred. No. 1.1e-19; Mismatches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ttttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 258
|||||
Db 81 ttttgccctcagctgaggttgctgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 19
AAC69120
ID AAC69120 standard; cDNA; 7864 BP.
AC AAC69120;
XX
XX 29-JAN-2001 (first entry)
XX Human ABC1 cholesterol transporter cDNA.
XX
XX Human ABC1 cholesterol transporter; chromosome 9q31;
XX ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
XX Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
XX cardiovascular disease; coronary artery disease; coronary restenosis;
XX cerebrovascular disease; peripheral vascular disease;
XX Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
XX X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
XX prognosis; prophylaxis; drug screening; transgenic animal; ss.
XX Homo sapiens.
XX OS
XX WO200055318-A2.
XX PN
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-IB00532.

PR 15-MAR-1999; 99US-0124702.
 PR 08-JUN-1999; 99US-0138048.
 PR 17-JUN-1999; 99US-0139600.
 PR 01-SEP-1999; 99US-0151977.
 XX
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 PA (XENO-) XENON BIORESEARCH INC.
 XX
 PI Hayden MR, Wilson AR, Pimstone SN;
 XX
 DR WPI: 2000-587528/55.
 DR P-PSDB; AAB38082.
 XX
 PT New ABC1 polypeptide is useful for treating diseases associated with
 PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
 PT disease and cancer -
 XX
 PS Claim 13; Page 157-160; 229pp; English.
 XX
 CC The invention relates to the human ABC1 cholesterol transporter protein
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of
 CC proteins, and plays a crucial role in cholesterol transport, particularly
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
 CC located on chromosome 9q31, and mutations in this gene are associated
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
 CC are distinguishable in that TD is an autosomal recessive disorder, while
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
 CC cholesterol") in the blood correlate with a high risk of cardiovascular
 CC disease, particularly coronary artery disease, but also cerebrovascular
 CC disease, coronary restenosis, and peripheral vascular disease.
 CC Conversely, a high level of HDL has protective effects against
 CC cardiovascular disease. The invention provides genetic constructs and
 CC transgenic cells and non-human animals comprising human ABC1 nucleic
 CC acids, and methods of gene therapy for the treatment or prevention of
 CC cardiovascular disease comprising the administration of an expression
 CC vector encoding ABC1 or an active fragment thereof. The invention also
 CC encompasses compounds which mimic ABC1 activity, compounds which
 CC stimulate ABC1 expression and methods of screening for such compounds.
 CC It further relates to methods for determining whether a patient has an
 CC increased risk for cardiovascular disease due to polymorphisms in the
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
 CC or prevent cardiovascular disease, especially coronary artery disease,
 CC cerebrovascular disease, coronary restenosis or peripheral vascular
 CC disease. They may also be used in the treatment of diseases associated
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
 CC The invention specifically excludes proteins with the exact amino acid
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
 CC present sequence represents cDNA encoding the human ABC1 cholesterol
 CC transporter.
 XX
 SQ Sequence 7864 BP: 2014 A; 1860 C; 2011 G; 1978 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7864;
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 199 tgttgccagctgaggttgctgtggaagaacctcacttctcagaagaaca 258
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81 tgttgccagctgaggttgctgtggaagaacctcacttctcagaagaaca 140
 ||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 20
 AAC69385
 ID AAC69385 standard; cDNA; 7864 BP.
 XX
 AC AAC69385;
 XX

DT 29-JAN-2001 (first entry)
 XX Human ABC1 cholesterol transporter TD-1 mutant cDNA (T4503C).
 DE
 XX Human ABC1 cholesterol transporter; chromosome 9q31;
 KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
 KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
 KW cardiovascular disease; coronary artery disease; coronary restenosis;
 KW cerebrovascular disease; peripheral vascular disease;
 KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
 KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
 KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.
 XX Homo sapiens.
 OS
 XX W0200055318-A2.
 PN
 XX 21-SEP-2000.
 PD
 XX 15-MAR-2000; 2000WO-IB00532.
 PF
 XX 15-MAR-1999; 99US-0124702.
 PR 08-JUN-1999; 99US-0138048.
 PR 17-JUN-1999; 99US-0139600.
 PR 01-SEP-1999; 99US-0151977.
 XX
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 PA (XENO-) XENON BIORESEARCH INC.
 XX
 PI Hayden MR, Wilson AR, Pimstone SN;
 XX
 DR WPI: 2000-587528/55.
 DR P-PSDB; AAB38104.
 XX
 PT New ABC1 polypeptide is useful for treating diseases associated with
 PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
 PT disease and cancer -
 XX
 PS Examples; Page -: 229pp; English.
 XX
 CC The invention relates to the human ABC1 cholesterol transporter protein
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of
 CC proteins, and plays a crucial role in cholesterol transport, particularly
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
 CC located on chromosome 9q31, and mutations in this gene are associated
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
 CC are distinguishable in that TD is an autosomal recessive disorder, while
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
 CC cholesterol") in the blood correlate with a high risk of cardiovascular
 CC disease, particularly coronary artery disease, but also cerebrovascular
 CC disease, coronary restenosis, and peripheral vascular disease.
 CC Conversely, a high level of HDL has protective effects against
 CC cardiovascular disease. The invention provides genetic constructs and
 CC transgenic cells and non-human animals comprising human ABC1 nucleic
 CC acids, and methods of gene therapy for the treatment or prevention of
 CC cardiovascular disease comprising the administration of an expression
 CC vector encoding ABC1 or an active fragment thereof. The invention also
 CC encompasses compounds which mimic ABC1 activity, compounds which
 CC stimulate ABC1 expression and methods of screening for such compounds.
 CC It further relates to methods for determining whether a patient has an
 CC increased risk for cardiovascular disease due to polymorphisms in the
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
 CC or prevent cardiovascular disease, especially coronary artery disease,
 CC cerebrovascular disease, coronary restenosis or peripheral vascular
 CC disease. They may also be used in the treatment of diseases associated
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
 CC The invention specifically excludes proteins with the exact amino acid
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
 CC present sequence represents cDNA encoding the human ABC1 cholesterol
 CC transporter.
 XX

CC present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC transporter associated with an altered cholesterol level and therefore an
CC altered risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 cDNA shown on pages 157-160.

XX Sequence 7864 BP; 2014 A; 1861 C; 2011 G; 1977 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7864;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggtgtgctgtggaagaacctcactttcagaagaagacaaaca 258
Db 81 tgttgccctcagctgaggtgtgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 21

AAC69386

ID AAC69386 standard; cDNA; 7864 BP.

XX AAC69386;

DT 29-JAN-2001 (first entry)

Human ABC1 cholesterol transporter TD-2 mutant cDNA (A1864G).

XX Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.

XX Homo sapiens.

OS WO200055318-A2.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-IB00532.

XX 15-MAR-1999; 99US-0124702.

PR 08-JUN-1999; 99US-0138048.

PR 17-JUN-1999; 99US-0139600.

PR 01-SEP-1999; 99US-0151977.

XX (UYBR-) UNIV BRITISH COLUMBIA.

FA (XENO-) XENON BIORESEARCH INC.

XX Hayden MR, Willson AR, Pimstone SN;

XX WPT; 2000-587528/55.

DR P-FSDB; AAB38105.

XX New ABC1 polypeptide is useful for treating diseases associated with

ABCI biological activity, e.g. Alzheimer's disease, Huntington's

disease and cancer -

XX Examples; Page -; 229pp; English.

XX The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases

are distinguishable in that TD is an autosomal recessive disorder, while
FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
cholesterol") in the blood correlate with a high risk of cardiovascular
disease, particularly coronary artery disease, but also cerebrovascular
disease, coronary restenosis, and peripheral vascular disease.
Conversely, a high level of HDL has protective effects against
cardiovascular disease. The invention provides genetic constructs and
transgenic cells and non-human animals comprising human ABC1 nucleic
acids, and methods of gene therapy for the treatment or prevention of
cardiovascular disease comprising the administration of an expression
vector encoding ABC1 or an active fragment thereof. The invention also
encompasses compounds which mimic ABC1 activity, compounds which
stimulate ABC1 expression and methods of screening for such compounds.
It further relates to methods for determining whether a patient has an
increased risk for cardiovascular disease due to polymorphisms in the
ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
or prevent cardiovascular disease, especially coronary artery disease,
cerebrovascular disease, coronary restenosis or peripheral vascular
disease. They may also be used in the treatment of diseases associated
with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
The invention specifically excludes proteins with the exact amino acid
sequences of GenBank Accession No: CAA10005.1 and X75926.1. The
acid with the exact sequence as GenBank Accession No: AJ012376.1. The
present sequence represents cDNA encoding a mutant human ABC1 cholesterol
transporter associated with an altered cholesterol level and therefore an
altered risk of cardiovascular disease.

CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 cDNA shown on pages 157-160.

XX Sequence 7864 BP; 2013 A; 1860 C; 2012 G; 1978 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7864;

Best Local Similarity 100.0%; Pred. No. 1.1e-19;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggtgtgctgtggaagaacctcactttcagaagaagacaaaca 258

Db 81 tgttgccctcagctgaggtgtgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 22

AAC69389

ID AAC69389 standard; cDNA; 7864 BP.

XX AAC69389;

DT 29-JAN-2001 (first entry)

Human ABC1 cholesterol transporter FHA-2 mutant cDNA (C6504T).

XX Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.

OS Homo sapiens.

OS WO200055318-A2.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-IB00532.

PR 08-JUN-1999; 99US-0124702.

PR 17-JUN-1999; 99US-0138048.

PR 01-SEP-1999; 99US-0139600.

PR 01-SEP-1999; 99US-0151977.

XX (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
XX WPI: 2000-587528/55.
DR P-PSDB; AAB38108.
XX
XX New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
XX Examples; Page -: 229pp; English.
XX
XX The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents cDNA encoding a mutant human ABC1 cholesterol
CC transporter associated with an altered cholesterol level and therefore an
CC altered risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 cDNA shown on pages 157-160.
XX
XX Sequence 7864 BP; 2014 A; 1859 C; 2011 G; 1979 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7864;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 tttgtgctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258
Db 81 tttgtgctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 140
|||||

RESULT 23
AAS06120
ID AAS06120 standard; cDNA; 9741 BP.
XX
XX AAS06120;
XX
XX 12-SEP-2001 (first entry)
DT

XX Human ABC1 DNA sequence #1.
DE
XX
KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW cardiovascular; neurological; Tangier disease; LCAT deficiency;
KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 185..6967
FT /*tag= a
FT /product= "Human ABC1 protein"
XX
PN WO200130848-A2.
XX
XX 03-MAY-2001.
XX
XX 26-OCT-2000; 2000WO-EP10886.
XX
XX 26-OCT-1999; 99EP-0402668.
XX
XX 01-MAR-2000; 2000US-0186260.
XX
XX (AVET) AVENTIS PHARMA SA.
XX
XX Denefle P, Rosier-Montus M, Arnould-Reguinne L, Prades C, Naudin L;
PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;
PI Dean M;
XX
XX WPI: 2001-316327/33.
XX
XX P-PSDB; AAU02176.
XX
XX New human ABC1 nucleic acids and polypeptides for treating
PT atherosclerosis, malaria and diabetes -
XX
XX Claim 1; Page 204-208; 368pp; English.
XX
XX The sequence represents the coding sequence #1 of human ABC1. The
CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,
CC and polypeptides and vectors are useful for the prevention of
CC atherosclerosis, in a subject affected by a dysfunction in the reverse
CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC useful for screening for an active ingredient for the prevention or
CC treatment of a disease resulting from dysfunction in the reverse
CC transport of cholesterol. The nucleic acids and polypeptides are also
CC useful for treating and preventing cardiovascular and neurological
CC pathologies, and other diseases e.g. Tangier disease, lecithin-
CC cholesterol (LCAT) deficiency, malaria and diabetes.
XX
XX Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;

Query Match 16.8%; Score 60; DB 22; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 tttgtgctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258
Db 191 tttgtgctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 250
|||||

RESULT 24
AAS06121
ID AAS06121 standard; cDNA; 9854 BP.
XX
XX AAS06121;
XX
XX 12-SEP-2001 (first entry)
DT
XX Human ABC1 DNA sequence #2.
DE
XX Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW cardiovascular; neurological; Tangier disease; LCAT deficiency;

KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 298..7078

XX FT /*tag= a

XX FT /product= "Human ABC1 protein"

XX WO200130848-A2.

XX PD 03-MAY-2001.

XX PF 26-OCT-2000; 2000WO-EP10886.

XX PR 26-OCT-1999; 99EP-0402668.

XX PR 01-MAR-2000; 2000US-0186260.

XX PA (AVET) AVENTIS PHARMA SA.

PI Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;

PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;

PI Dean M;

DR WPI: 2001-316327/33.

DR P-PSDB; AAU02176.

XX New human ABC1 nucleic acids and polypeptides for treating

PT atherosclerosis, malaria and diabetes -

XX Claim 1; Page 209-213; 368pp; English.

CC The sequence represents the coding sequence #2 of human ABC1. The
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,
 CC and polypeptides and vectors are useful for the prevention of
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is
 CC useful for screening for an active ingredient for the prevention or
 CC treatment of a disease resulting from dysfunction in the reverse
 CC transport of cholesterol. The nucleic acids and polypeptides are also
 CC useful for treating and preventing cardiovascular and neurological
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-
 CC cholesterol (LCAT) deficiency, malaria and diabetes.

XX Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;

Query Match 16.8%; Score 60; DB 22; Length 9854;

Best Local Similarity 100.0%; Pred. No. 1.1e-19;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 tttggcctcagctgaggttgctgtgtggaagaacctcactttcagaagaagacaaca 258

|||||

Db 304 tttggcctcagctgaggttgctgtgtggaagaacctcactttcagaagaagacaaca 363

RESULT 25

AAH07432

ID AAH07432 standard; cDNA; 736 BP.

XX AC AAH07432;

XX 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:4267.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

PN EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 14.3%; Score 51; DB 22; Length 736;

Best Local Similarity 100.0%; Pred. No. 2.9e-15;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 208 cagctgaggttgctgtgtggaagaacctcactttcagaagaagacaaca 258

|||||

Db 329 cagctgaggttgctgtgtggaagaacctcactttcagaagaagacaaca 379

RESULT 26

AAH18606

ID AAH18606 standard; cDNA; 1556 BP.

XX AC AAH18606;

XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:18808.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

PN EP1074617-A2.

XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 17-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI; 2001-318749/34.
 XX PR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.
 XX CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX SQ Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

Query Match 14.3%; Score 51; DB 22; Length 1556;
 Best Local Similarity 100.0%; Pred. No. 2.9e-15;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 208 cagctgaggttgcctgtggaagaacctcatttcagaagaagacaaca 258
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 329 cagctgaggttgcctgtggaagaacctcatttcagaagaagacaaca 379

RESULT 27
 AAF93084
 ID AAF93084 standard; DNA; 37 BP.
 XX AC AAF93084;
 XX DT 17-MAY-2001 (first entry)
 XX DE ABC1 polymorphism RFLP oligonucleotide #45.
 XX KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
 XX OS Homo sapiens.

XX WO200115676-A2.
 XX PD 08-MAR-2001.
 XX PR 01-SEP-2000; 2000WO-IB01492.
 XX PR 01-SEP-1999; 99US-0151977.
 XX PR 15-MAR-2000; 2000US-0526193.
 XX PR 23-JUN-2000; 2000US-0213958.
 XX PA (UYBR-) UNIV BRITISH COLUMBIA.
 XX PA (XENO-) XENON GENETICS INC.
 XX PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
 XX DR WPI; 2001-244356/25.
 XX PR Treating a lower than normal high density lipoprotein-cholesterol
 PT (HDL-C) level, a higher than normal triglyceride level, or a
 PT cardiovascular disease, by administering a compound that modulates LXR-
 PT or RXR-mediated transcriptional activity -
 XX PS Disclosure; Fig 17; 317pp; English.
 XX CC The present invention relates to a method for treating a patient
 CC diagnosed as having a lower than normal high density
 CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
 CC triglyceride level, or a cardiovascular disease, involving
 CC administering a compound that modulates LXR- or RXR-mediated
 CC transcriptional activity or ABC1 expression or activity.
 CC The LXR gene product may be used in an assay to identify
 CC compounds useful for the treatment of a disease or condition selected a
 CC lower than normal HDL cholesterol level, a higher than normal
 CC triglyceride level, and a cardiovascular disease.
 XX SQ Sequence 37 BP; 4 A; 17 C; 11 G; 5 T; 0 other;
 Query Match 9.8%; Score 35; DB 22; Length 37;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 138 gcgcgtccttcagggtcccgagccacacgctg 172
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1 gcgcgtccttcagggtcccgagccacacgctg 35
 RESULT 28
 AAF93082
 ID AAF93082 standard; DNA; 38 BP.
 XX AC AAF93082;
 XX DT 17-MAY-2001 (first entry)
 XX DE ABC1 polymorphism RFLP oligonucleotide #43.
 XX KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
 XX OS Homo sapiens.
 XX PA WO200115676-A2.
 XX PD 08-MAR-2001.
 XX PR 01-SEP-2000; 2000WO-IB01492.
 XX PR 01-SEP-1999; 99US-0151977.
 XX PR 15-MAR-2000; 2000US-0526193.
 XX PR 23-JUN-2000; 2000US-0213958.
 XX PA (UYBR-) UNIV BRITISH COLUMBIA.

PA (XENO-) XENON GENETICS INC.
XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX WPI; 2001-244356/25.
XX
XX Treating a lower than normal high density lipoprotein-cholesterol
XX (HDL-C) level, a higher than normal triglyceride level, or a
XX cardiovascular disease, by administering a compound that modulates LXR-
XX or RXR-mediated transcriptional activity -
XX
XX Disclosure; Fig 17; 317pp; English.
XX
XX The present invention relates to a method for treating a patient
XX diagnosed as having a lower than normal high density
XX lipoprotein-cholesterol (HDL-C) level, a higher than normal
XX triglyceride level, or a cardiovascular disease, involving
XX administering a compound that modulates LXR- or RXR-mediated
XX transcriptional activity or ABCI expression or activity.
XX The LXR gene product may be used in an assay to identify
XX compounds useful for the treatment of a disease or condition selected a
XX lower than normal HDL cholesterol level, a higher than normal
XX triglyceride level, and a cardiovascular disease.
XX
XX Sequence 38 BP; 6 A; 15 C; 5 G; 12 T; 0 other;
SQ

Query Match 9.8%; Score 35; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 ccctctgcttattcttcagtttaagtaccagcc 112
|||||
Db 1 cccctctgcttattcttcagtttaagtaccagcc 35

RESULT 29
AAH26495
ID AAH26495 standard; DNA; 12425 BP.
XX
XX AAH26495;
XX
XX 12-NOV-2001 (first entry)
XX
XX Human low density lipoprotein binding protein 2 (LBP-2) gene.
DE
XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;
XX atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis;
XX vaccine; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 2832..5153
FT CDS
FT /*tag= a
FT /*note= "includes introns"
FT exon 2832..3785
FT /*tag= b
FT intron 3786..4207
FT /*tag= c
FT exon 4208..4502
FT /*tag= d
FT intron 4503..4593
FT /*tag= e
FT exon 4594..4694
FT /*tag= f
FT intron 4695..4787
FT /*tag= g
FT exon 4788..4899
FT /*tag= h
FT intron 4900..4994
FT /*tag= i
FT exon 4995..5153

FT /*tag= j
XX WO200164874-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 28-FEB-2001; 2001WO-US06356.
PF
XX
XX 02-MAR-2000; 2000US-0517849.
PR
XX 14-JUL-2000; 2000US-0616289.
PR
XX
XX (BOST-) BOSTON HEART FOUND INC.
PA
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
PI
XX WPI; 2001-565505/63.
DR
XX P-PSDB; AAH82806.
DR
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
XX diagnosing and/or identifying therapeutic agents for atherosclerosis -
XX
XX Example 4; Fig 23; 143pp; English.
PS
XX
XX The present sequence is that of genomic DNA encoding novel human
XX low density lipoprotein binding protein 2 (LBP-2, see AAH82806).
XX The DNA was isolated from a human genomic library by screening with
XX LBP-2 cDNA (see AAH26494). The open reading frame spans 5 exons.
XX Human LBP-2 nucleic acids are among claimed polynucleotides of the
XX invention that encode novel polypeptides, termed LBPs, capable of
XX binding to native and methylated LDL. Also claimed are isolated
XX LBP polypeptides, and biologically active fragments and analogues
XX of them, as well as expression vectors, cells and methods of
XX producing the LBPs. Methods of determining if an animal is at risk
XX for atherosclerosis, and methods for evaluating an agent for use in
XX treating atherosclerosis, and methods for treating a cell having an
XX abnormality in structure or metabolism of LBP are claimed.
XX pharmaceutical compositions comprising an LBP polypeptide or
XX nucleic acid, and vaccine compositions, are also claimed.
XX
XX Sequence 12425 BP; 2563 A; 3702 C; 3566 G; 2594 T; 0 other;
SQ

Query Match 6.7%; Score 24; DB 22; Length 12425;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggaggaggaggagg 349
|||||
Db 2395 ggaggaggaggaggaggaggaggaggaggagg 2418

RESULT 30
AAZ39175
ID AAZ39175 standard; DNA; 30 BP.
XX
XX AAZ39175;
XX
XX 06-MAR-2000 (first entry)
DT
XX
XX Human BMP-4 5' upstream gene sequence with exon 1-3 PCR primer #1.
DE
XX
XX Human; bone morphogenetic protein 4; BMP-4; promoter; cartilage;
XX bone disease; osteohyperplasia; osteometastasis; orthopaedic surgery;
XX osteoarthritis; arthrosteitis; tumour dissection; bone reconstruction;
XX spinal fusion; vertebral canal enlargement; congenital cartilage disease;
XX dysosteogenesis; achondroplasia; palatoschisis; mandible reconstruction;
XX residual ridge construction; osteoporosis; morphogenesis; hyperplasia;
XX PCR primer; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX WO957145-A2.
PN


```

    reduce
    me

    Gaps    0;

    n;

```


CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1832 BP; 447 A; 439 C; 432 G; 484 T; 30 other;

Query Match 6.2%; Score 22; DB 23; Length 1832;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 328 agagaggaggaggaggaggag 349
 |||||
 Db 1321 AGGAGGGAGGGAGGAGGAG 1300

RESULT 42

AAT78942

ID AAT78942 standard; cDNA; 1944 BP.

XX AC AAT78942;

XX DT 22-JAN-1998 (first entry)

XX DE Human bone morphogenic protein (BMP) 2B cDNA.

XX DE bone morphogenic protein; BMP 2B; cartilage; periodontal disease;

KW tissue repair; osteoporosis; treatment; oligonucleotide probe; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT 5'UTR 8..402

FT /*tag= a

FT CDS 403..1629

FT /*tag= b

FT /product= Bone-morphogenic_protein_2B

FT mat_peptide 1333..1626

FT /*tag= c

FT /note= "BMP-2B contains at least this part of the

FT 3'UTR 1630..1938

FT /*tag= d

XX US5631142-A.

XX PD 20-MAY-1997.

XX PF 07-SEP-1993; 93US-0118363.

XX PR 11-JUL-1989; 89US-0378537.

XX PR 17-DEC-1986; 86US-0943332.

XX PR 20-MAR-1987; 87US-0028285.

XX PR 08-APR-1988; 88US-0179100.

XX PR 18-MAY-1992; 92US-0884353.

XX PR 07-SEP-1993; 93US-0118363.

XX PR 01-JUL-1986; 86US-0880776.

XX PA (GENY) GENETICS INST INC.

XX PI Rosen VA, Wang EA, Wozney JM;

XX XX WPI; 1997-288573/26.

XX DR P-PSDB; AAN24850.

XX XX

PT Production of human bone morphogenic protein 2A or 2B in cell
 PT culture - useful inducing bone or cartilage production, in wound
 PT healing and tissue repair
 XX
 FS Example 5; Fig 3; 22pp; English.
 XX
 CC This cDNA sequence encodes the human bone morphogenic protein (BMP) 2B.
 CC Oligonucleotide probes were synthesised based on a partially sequenced
 CC BMP-2B protein isolated from ground bovine powder. The probes were used
 CC to screen a bovine liver DNA library to obtain the BMP-2B encoding DNA
 CC sequence. The DNA was used to screen a U-2 OS human cell line cDNA
 CC library to obtain this human BMP-2B cDNA. BMPs can be used to induce
 CC bone and cartilage formation, and in wound healing and tissue repair.
 CC They can also be used for treating periodontal disease or osteoporosis.
 XX
 SQ Sequence 1944 BP; 488 A; 535 C; 510 G; 411 T; 0 other;

Query Match 6.2%; Score 22; DB 18; Length 1944;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggagggaggaggaggaggaggag 347

|||||

Db 15 ggagggaggaggaggaggaggag 36

RESULT 43

AAN80634

ID AAN80634 standard; DNA; 1954 BP.

XX AC AAN80634;

XX XX 08-OCT-1990 (first entry)

XX DE Human Bone Morphogenic Protein-2 class II cDNA from U2OS-3.

XX DE bone morphogenic protein; hBMP-2 class II; probes; cartilage formation;

KW bone formation; osteogenic cpds; prodontal disease; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 403..1623

FT /*tag= a

FT /product=hBMP-2 class II

XX WO8800205-A.

XX PD 14-JAN-1988.

XX XX 30-JUN-1987; 87WO-US01537.

XX XX 26-MAR-1987; 87US-0031346.

XX PA (GENE-) GENETICS INST INC.

XX PI Wozney JM, Rosen VA;

XX XX WPI; 1988-021565/03.

XX DR P-PSDB; AAP80620.

XX XX Bone morphogenic proteins - obtd. using recombinant DNA and used

XX PT for inducing cartilage and bone formation.

XX PS Disclosure; ; 7pp; English.

XX XX The HindIII-SacI bovine genomic bBMP-2 fragment described in AAN80627

XX CC is subcloned into M13, labelled and used as probe to screen

XX CC polyadenylated RNAs from various cells and tissue sources.

XX CC Sequence analysis of the weakly hybridising recombinants hBMP-2 class II

XX CC (-BMP-4) indicated that they are quite homologous with the sequence

XX CC given in AAN80622 at the end of their coding regions, but less so in

Sequence 1954 BP; 492 A; 536 C; 511 G; 415 T; 0 other;
XX SQ

Qy 326 ggaggagggagggaaggga 347
 |||||
Dd 15 ggaggagggagggaaggga 36

Search completed: September 20, 2002, 06:07:38
Job time: 10342 sec

[illegible]

Db 15 GGAGGAGGAGGAGGAAGGA 36

RESULT 7

US-08-449-700-6
; Sequence 6, Application US/08449700
; Patent No. 5863758
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,700
; FILING DATE: 21-FEB-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 810,560
; FILING DATE: 20-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 827,052
; FILING DATE: 28-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 660,162
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,988
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,849
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 616,374
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,024
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 599,543
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 579,865
; FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 569,920
; FILING DATE: 20-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 483,913
; FILING DATE: 22-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 422,613
; FILING DATE: 17-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 315,342
; FILING DATE: 23-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 232,630
; FILING DATE: 15-AUG-1988

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 179,460
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-001CP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1788 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; TISSUE TYPE: HIPPOCAMPUS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..1626
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
; OTHER INFORMATION: /product= "CBMP2B"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /note= "CBMP2B (CDNA)"
US-08-449-700-6

Query Match 6.2%; Score 22; DB 2; Length 1788;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggagggaggaggaggagga 347
|||||
Db 15 GGAGGAGGAGGAGGAAGGA 36

RESULT 8
US-08-449-699A-6
; Sequence 6, Application US/08449699A
; Patent No. 5958441
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: ANTIBODIES TO OSTEOGENIC PROTEINS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,699A
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/147,023
; FILING DATE: 01-NOV-1993

Query Match 6.28; Score 22; DB 2; Length 1954;

```

RESULT 12
US-08-750-222A-3
; Sequence 3, Application US/08750222A
; Patent No. 6034061
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Song, Jeffrey
; APPLICANT: Thies, Scott
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: Legal Affairs - 87 Cambridge
; CITY: Cambridge
; STATE: MA
; COUNTRY: US
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Ver
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,222A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,353
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Karpinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/POCKET NUMBER: GI 5186B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Osteosarcoma Cell Line
; CELL LINE: U-2OS
; IMMEDIATE SOURCE:
; LIBRARY: U2OS cDNA in Lambda gt10
; CLONE: Lambda U2OS-3
; POSITION IN GENOME:
; UNITS: bp
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..1629
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1279..1626
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: 9..1934

```


Query Match 6.2%; Score 22; DB 3; Length 1954;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggaggaggaggaggaggaggga 347
Db 15 GGAGGAGGAGGAGGAGGAGGA 36

RESULT 15
US-08-469-411-3
; Sequence 3, Application US/08469411
; Patent No. 6190880
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; Wolfman, Neil M.
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; Heterodimers, Compositions and Methods of Use.
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,411
; FILING DATE: 06-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI-5192B-CON
; TELEPHONE: 617-498-8622
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..1626
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-469-411-3

Query Match 6.2%; Score 22; DB 4; Length 1954;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggaggaggaggaggaggaggga 347
Db 15 GGAGGAGGAGGAGGAGGAGGA 36

RESULT 16
US-08-925-779-5
; Sequence 5, Application US/08925779
; Patent No. 6245889
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth A.
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.

; TITLE OF INVENTION: NO. 6245889el BMP Products
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
; STREET: 87 CAMBRIDGEPARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,779
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/721,847
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: 5160C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; IMMEDIATE SOURCE:
; LIBRARY: U2OS CDNA in Lambda gt10
; CLONE: Lambda U2OS-3
; POSITION IN GENOME:
; UNITS: bp
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..1629
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1279..1629
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: 9..1934
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-925-779-5

Query Match 6.2%; Score 22; DB 4; Length 1954;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggaggaggaggaggaggaggga 347
Db 15 GGAGGAGGAGGAGGAGGAGGA 36

RESULT 17
US-08-254-353A-3
; Sequence 3, Application US/08254353A
; Patent No. 6287816
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.

Query Match 6.2%; Score 22; DB 5; Length 1954;
Best Local Similarity 100.0%; Pred. No. 0.053;

QY 326 ggaggaggaggaggaggga 347
|||
Db 15 GCAGGAGGAGGGGCAAGGA 36

```

RESULT 19
PCT-US95-07084-3
; Sequence 3, Application PC/TUS9507084
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
;

```


Db 340 GGAGGGAGGAGGAAGGAAG 320

RESULT 22

US-08-332-766A-17/C
; Sequence 17, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332.766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-332-766A-17

Query Match 5.9%; Score 21; DB 2; Length 591;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggagggaggagggaaggaag 349
|||||
Db 503 GGAGGGAGGAGGAAGGAAG 483

RESULT 23

US-08-332-766A-42
; Sequence 42, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.

; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332.766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-332-766A-42

Query Match 5.6%; Score 20; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agagggaggagggaagga 347
|||||
Db 196 AGGAGGGAGGGAAGGA 215

RESULT 24

US-09-268-992-7
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freilmer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match 5.6%; Score 20; DB 4; Length 72604;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 ttccccctctgtttatct 94
|||||
Db 14772 ttccccctctgtttatct 14791

RESULT 25
US-08-332-766A-27
; Sequence 27, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-27

Query Match 5.3%; Score 19; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggagggaagga 347
|||||
Db 152 GGAGGAGGAGGGAAGGA 170

RESULT 26
US-08-332-766A-22
; Sequence 22, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-22

Query Match 5.3%; Score 19; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 agggaggagggaaggaag 349
|||||
Db 205 AGGAGGAGGGAAGGAAG 223

RESULT 27
US-08-713-000-9
; Sequence 9, Application US/08713000
; Patent No. 5850020
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Grierson, Alastair
; TITLE OF INVENTION: MATERIALS AND METHODS FOR THE
; MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,000
; FILING DATE:
; CLASSIFICATION: 435

```

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 gctggcgctgctggctgag 187
      |||||
Db 261 GCTGGCGTCTGGCTGAG 279

RESULT 29
US-09-211-710-9
; Sequence 9, Application US/09211710A
; Patent No. 6204434
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkaia, Ilkka
; APPLICANT: Grierson, Alastair
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c3
; CURRENT APPLICATION NUMBER: US/09/211,710A

```

```

: TITLE OF INVENTION: Materials and Methods for the
:
: FILE REFERENCE: 11000.1003c3
:
: CURRENT APPLICATION NUMBER: US/09/211,710A
:
: CURRENT FILING DATE: 1998-12-14
:
: NUMBER OF SEQ ID NOS: 15
:
: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO. 9
:

```

/ TYPE: DNA
 ; ORGANISM: Pinus radiata
 US-09-211-710-9

```

Query Match          5.3%; Score 19; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 gctggggcgtgctggctgag 187
      |||||||
Db 261 gctgggcgtgctggctgag 279

RESULT 30
US-08-975-316-45
; Sequence 45, Application US/08975316
Date of Seq. 5052486

```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

; REFERENCE/DOCKET NUMBER: 11000/1003C1
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
;

```

; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-975-316-45

Query Match 5.3%; Score 19; DB 2; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 gctggcgctgctgctgag 187
|||||
Db 261 GCTGGCGTCTGCTGAG 279

RESULT 31
US-08-705-771-8
; Sequence 8, Application US/08705771
; Patent No. 6054289
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; APPLICANT: Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; TITLE OF INVENTION: Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,771
; FILING DATE: August 30, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1344 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-705-771-8

Query Match 5.3%; Score 19; DB 3; Length 1344;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaagagagagagagaa 344
|||||
Db 959 GGAGGAGGAGGAGGAGAA 977

RESULT 32
US-09-370-807-7/c
; Sequence 7, Application US/09370807
; Patent No. 6297034
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/370,807
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/096,225
; EARLIER FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-370-807-7

Query Match 5.0%; Score 18; DB 4; Length 2407;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 gcttgacgacataactg 67
|||||
Db 513 GCTTGCAGCAATAACTG 496

RESULT 33
US-08-765-662-13
; Sequence 13, Application US/08765662
; Patent No. 5929213
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-12
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,662
; FILING DATE: 28-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08745
; FILING DATE: 12-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/042W01 (FD-3830)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2419 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 218...1267
; OTHER INFORMATION:
US-08-765-662-13

Query Match 5.0%; Score 18; DB 2; Length 2419;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggagg 346
|||||
Db 1563 GGAGGGAGGGGGAAGG 1580

RESULT 34

PCT-US95-08745-13
; Sequence 13, Application PC/TUS9508745
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-12
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08745
; FILING DATE: 12-JUL-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/042W01 (FD-3830)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2419 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 218...1267
; OTHER INFORMATION:
PCT-US95-08745-13

Query Match 5.0%; Score 18; DB 5; Length 2419;
Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 329 ggaggaggaggaggagg 346
|||||
Db 1563 GGAGGGAGGGGGAAGG 1580

RESULT 35

US-09-173-914-1
; Sequence 1, Application US/09173914
; Patent No. 6171857
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Eric
; TITLE OF INVENTION: A No. 6171857a1 Leucine zipper, KARP-1 and
; FILE REFERENCE: 80877/7017/HK
; CURRENT APPLICATION NUMBER: US/09/173,914
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 60/064,557
; EARLIER FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 6078
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (178)...(178)
; NAME/KEY: unsure
; LOCATION: (230)...(230)
; NAME/KEY: unsure
; LOCATION: (232)...(232)
; NAME/KEY: unsure
; LOCATION: (234)...(234)
; NAME/KEY: unsure
; LOCATION: (453)...(453)
; NAME/KEY: unsure
; LOCATION: (473)...(473)
; NAME/KEY: unsure
; LOCATION: (610)...(610)
; NAME/KEY: unsure
; LOCATION: (612)...(612)
; NAME/KEY: unsure
; LOCATION: (2175)...(2175)
; NAME/KEY: unsure
; LOCATION: (1014)...(1014)
US-09-173-914-1

Query Match 5.0%; Score 18; DB 4; Length 6078;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 gagggaggaggaggagg 347
|||||
Db 204 gagggaggaggaggagg 221

RESULT 36

US-08-483-376-1/c
; Sequence 1, Application US/08483376
; Patent No. 5955330
; GENERAL INFORMATION:
; APPLICANT: Vasil, Vimla
; APPLICANT: Clancy, Maureen A.
; APPLICANT: Ferl, Robert J.
; APPLICANT: Vasil, Indra K.
; APPLICANT: Hannah, L. C.
; TITLE OF INVENTION: No. 5955330e1 Means for Enhancing Gene
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
US-08-483-376-1/c

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/483,376
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/418,540
FILING DATE: 07-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,115
FILING DATE: 04-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/830,956
FILING DATE: 05-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/353,854
FILING DATE: 18-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 10-94B
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6386 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Zea mays
STRAIN: Black Sweet
FEATURE:
NAME/KEY: exon
LOCATION: 131..182
FEATURE:
NAME/KEY: exon
LOCATION: 1211..1324
FEATURE:
NAME/KEY: exon
LOCATION: 1828..1948
FEATURE:
NAME/KEY: exon
LOCATION: 2041..2187
FEATURE:
NAME/KEY: exon
LOCATION: 2269..2460
FEATURE:
NAME/KEY: exon
LOCATION: 2605..2728
FEATURE:
NAME/KEY: exon
LOCATION: 2822..3038
FEATURE:
NAME/KEY: exon
LOCATION: 3256..3351
FEATURE:
NAME/KEY: exon
LOCATION: 3447..3620

FEATURE:
NAME/KEY: exon
LOCATION: 3702..3818
FEATURE:
NAME/KEY: exon
LOCATION: 3912..4078
FEATURE:
NAME/KEY: exon
LOCATION: 4158..4381
FEATURE:
NAME/KEY: exon
LOCATION: 4517..4835
FEATURE:
NAME/KEY: exon
LOCATION: 4768..5212
FEATURE:
NAME/KEY: exon
LOCATION: 5372..5510
FEATURE:
NAME/KEY: exon
LOCATION: 5636..5917
US-08-483-376-1
Query Match 5.0%; Score 18; DB 2: Length 6386;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 325 tggagggaggaggaggagg 342
|||||
Db 151 TGGAGGAGGAGGAGGAGG 134
RESULT 37
US-08-814-095-7/c
Sequence 7, Application US/08814095
Patent No. 6025183
GENERAL INFORMATION:
APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Shani, Moshe
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 No. 6025183thwestern Highway, Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 35060 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Cosmid including ACHE"
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 7q22
FEATURE:
NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHE Promotor"
OTHER INFORMATION: /standard_name= "ACHE Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 27005..27274
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 27255..28007
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: terminator
LOCATION: 27385..27387
FEATURE:
NAME/KEY: exon
LOCATION: 28008..28129
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: 34528..34895
OTHER INFORMATION: /function= "arsenite resistance"
OTHER INFORMATION: /gene= "ars"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 34092..34358
```

```
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: complement (33779..33963)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: complement (33493..33591)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: complement (33297..33408)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: complement (32959..33094)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: complement (32569..32628)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: complement (32386..32468)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: complement (31894..32080)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: complement (31363..31534)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: complement (31131..31284)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: complement (30816..31011)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: complement (30470..30626)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: complement (30187..30274)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: exon
LOCATION: complement (29945..30073)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: complement (29664..29856)
OTHER INFORMATION: /gene= "ARS"
OTHER INFORMATION: /number= 16
US-08-814-095-7
```

APPLICANT: CHENIVESSE, Xavier
APPLICANT: HENGLEIN, Berthold
APPLICANT: ZINDY, Fr d rique
TITLE OF INVENTION: NEW HUMAN CYCLIN A COMPOSITIONS AND A PROCESS FOR THEIR PRO
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Larson and Taylor
STREET: 727 Twenty-Third Street, South
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-5.25 inch, 500 Kb storage
COMPUTER: IBM PC/XT/AT or compatibles
OPERATING SYSTEM: MS-DOS version 3.0 or above
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,895
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/650, 805
FILING DATE: 06-FEB-1991
APPLICATION NUMBER: FR9001596
FILING DATE: 12-FEB-1990
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1634 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
ORGANISM: HOMO SAPIENS
FEATURE:
NAME/KEY: Human cyclin A
LOCATION: coding sequence from base 97 to base 1392,
LOCATION: coding for a protein of 432 amino acids.
US-08-460-895-1
Query Match 4.8%; Score 17; DB 2; Length 1634;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 56 cagcaataactgagc 72
Db 1165 CAGCAATAACTGATGCG 1149
RESULT 40
US-09-210-889-1/c
Sequence 1, Application US/09210889
Patent No. 6103887
GENERAL INFORMATION:
APPLICANT: BRECHOT, Christian
APPLICANT: WANG, Jian
APPLICANT: CHENIVESSE, Xavier
APPLICANT: HENGLEIN, Berthold
APPLICANT: ZINDY, Fr d rique
TITLE OF INVENTION: NEW HUMAN CYCLIN A COMPOSITIONS AND A
TITLE OF INVENTION: PROCESS FOR THEIR PRODUCTION, THE CORRESPONDING NUCLEOTIDE
TITLE OF INVENTION: A PROCESS AND AGENTS FOR CELL PROLIFERATION DETECTION OR DI
TITLE OF INVENTION: AND A PROCESS AND AGENTS FOR INHIBITING CELL PROLIFERATION.
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Larson and Taylor
STREET: 727 Twenty-Third Street, South
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202

Query Match 5.0%; Score 18; DB 3; Length 35060;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 330 gagggaggagggaagga 347
Db 23954 GAGGAGGAGGGAAGGA 23937
RESULT 38
US-08-332-766A-4/c
Sequence 4, Application US/08332766A
Patent No. 5843647
GENERAL INFORMATION:
APPLICANT: JEFFREYS, Alec J.
APPLICANT: ARMOUR, John
TITLE OF INVENTION: SIMPLE TANDEM REPEATS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,766A
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326052.9
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BIRD, Donald J.
REGISTRATION NUMBER: 25,323
REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-332-766A-4
Query Match 4.8%; Score 17; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 331 agggaggagggaagga 347
Db 102 AGGAGGAGGGAAGGA 86
RESULT 39
US-08-460-895-1/c
Sequence 1, Application US/08460895
Patent No. 5849508
GENERAL INFORMATION:
APPLICANT: BRECHOT, Christian
APPLICANT: WANG, Jian

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-5.25 inch, 500 kb storage
COMPUTER: IBM PC/XT/AT or compatibles
OPERATING SYSTEM: MS-DOS version 3.0 or above
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,889
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/650, 805
FILING DATE: 06-FEB-1991
APPLICATION NUMBER: FR9001596
FILING DATE: 12-FEB-1990
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1634 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: HOMO SAPIENS
FEATURE:
NAME/KEY: Human cyclin A
LOCATION: coding sequence from base 97 to base 1392,
LOCATION: coding for a protein of 432 amino acids.
US-09-210-889-1

Query Match 4.8%; Score 17; DB 3; Length 1634;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 cagcaataactgatgccc 72
|||||
DB 1165 CAGCAATAACTGATGCG 1149

RESULT 41
US-08-692-787-8/c
Sequence 8, Application US/08692787
Patent No. 5882864
GENERAL INFORMATION:
APPLICANT: An, Gang
APPLICANT: O'Hara, S. Mark
APPLICANT: Ralph, David
APPLICANT: Veltri, Robert
TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE
DISEASE
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/692,787
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Corder, Timothy S.
REGISTRATION NUMBER: 38,414
REFERENCE/DOCKET NUMBER: UROC-012
TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1649 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-692-787-8

Query Match 4.8%; Score 17; DB 2; Length 1649;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 cagcaataactgatgccc 72
|||||
DB 1176 CAGCAATAACTGATGCG 1160

RESULT 42
US-09-199-199-8/c
Sequence 8, Application US/09097199
Patent No. 6218529
GENERAL INFORMATION:
APPLICANT: An, Gang
APPLICANT: O'Hara, S. Mark
APPLICANT: Ralph, David
APPLICANT: Veltri, Robert
TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/097,199
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/692,787
FILING DATE: 31-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nakashima, Richard A.
REGISTRATION NUMBER: P-42,023
REFERENCE/DOCKET NUMBER: UROC:018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1649 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-097-199-8

Query Match 4.8%; Score 17; DB 4; Length 1649;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 cagcaataactgatgccc 72
|||||

Db 1176 CAGCAATACTGATGGC 1160

RESULT 43

US-09-078-294-4/c
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

Query Match 4.8%; Score 17; DB 4; Length 80246;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggagggaag 345
|||||
Db 21481 GGAGGGAGGGAGGAAG 21465

RESULT 44

US-09-078-294-3/c
; Sequence 3, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 80595
; TYPE: DNA
; ORGANISM: Nucleotide sequence of HC-contig
US-09-078-294-3

Query Match 4.8%; Score 17; DB 4; Length 80595;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggagggaag 345
|||||
Db 21743 GGAGGGAGGGAGGAAG 21727

RESULT 45

US-09-109-663-14
; Sequence 14, Application US/09109663
; Patent No. 6277981
; GENERAL INFORMATION:
; APPLICANT: Tu, Guang-Chou
; APPLICANT: Israel, Yedy
; TITLE OF INVENTION: AN IMPROVED METHOD FOR DESIGN AND SELECTION OF
; TITLE OF INVENTION: EFFICACIOUS ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 9855-301

; CURRENT APPLICATION NUMBER: US/09/109,663
; CURRENT FILING DATE: 1998-07-03
; EARLIER APPLICATION NUMBER: 60/051,705
; EARLIER FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Candidate
; OTHER INFORMATION: TNF(alpha) ASO
US-09-109-663-14

Query Match 4.5%; Score 16; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 gagggagggaaggaag 349
|||||
Db 3 gagggagggaaggaag 18

Search completed: September 20, 2002, 06:15:23
Job time: 11142 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 04:07:18 ; Search time 3900.56 Seconds
(without alignments)
1235.313 Million cell updates/sec

Title: US-09-846-456-2
Perfect score: 357
Sequence: 1 ttgaggtctcactgaaggg.....gagggaaggaactgtgttg 357

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hctc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	14.3	736	9	AU135588
2	39	10.9	535	10	BG384217
3	24	6.7	158	12	AQ508395
4	24	6.7	292	10	244377 HSC12B081 n
5	24	6.7	709	12	AQ528018
6	24	6.7	716	12	AG163359 Pan trogl
7	23	6.4	401	9	AT394181 tg67a04.x
8	23	6.4	827	12	B20638
9	22	6.2	282	9	BE050974
10	22	6.2	309	12	BH293527
11	22	6.2	371	10	BF592428
12	22	6.2	406	10	R11136
13	22	6.2	465	9	AT481405
14	22	6.2	478	9	BE050973
15	22	6.2	483	12	AZ654841
16	22	6.2	515	12	BH337839
17	22	6.2	523	10	BM310687

C 18	6.2	22	525	12	AZ647100
C 19	6.2	22	544	12	AQ415718
C 20	6.2	22	552	9	AV938157
C 21	6.2	22	569	12	AQ078910
C 22	6.2	22	572	12	AZ378099
C 23	6.2	22	591	12	AZ828464
C 24	6.2	22	615	12	AQ924975
C 25	6.2	22	623	12	B70930
C 26	6.2	22	673	10	BF984535
C 27	6.2	22	713	12	AZ742461
C 28	6.2	22	754	12	AQ983682
C 29	6.2	22	760	10	BE877424
C 30	6.2	22	907	10	BE617281
C 31	6.2	22	926	10	BG823588
C 32	6.2	22	1068	10	BG328313
C 33	5.9	21	43	12	AZ995124
C 34	5.9	21	51	12	AZ351915
C 35	5.9	21	52	12	AZ500262
C 36	5.9	21	66	12	AZ411857
C 37	5.9	21	90	12	AZ838528
C 38	5.9	21	120	12	AZ743191
C 39	5.9	21	145	9	AI053852
C 40	5.9	21	161	12	AZ091117
C 41	5.9	21	191	12	AZ457379
C 42	5.9	21	212	12	AZ656509
C 43	5.9	21	269	10	F11105
C 44	5.9	21	294	12	AZ513751
C 45	5.9	21	319	12	B50482

ALIGNMENTS

RESULT 1

AU135588
LOCUS AU135588 PLACEL Homo sapiens cDNA clone PLACE1002437 5', mRNA
sequence.
ACCESSION AU135588.1 GI:10996127
VERSION AU135588
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 736)
AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.

TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES
Location/Qualifiers
1..736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pME18SFL3"

source

BASE COUNT 163 a 199 c 199 g 170 t 5 others
ORIGIN

Query Match 14.3%; Score 51; DB 9; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.8e-14;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 cagctgaggttgcctgtctggaagaaccttcttcagagaagaaca 258
 |||||||
 Db 329 CAGCTGAGGTTGCTGCTGTGGAAGACCTCACTTTTCAGAGAAGACAAACA 379
 |||||||

RESULT 2
 BG384217
 LOCUS 303216 MARC LP1G Sus scrofa cDNA 5', mRNA linear EST 12-MAR-2001
 DEFINITION
 ACCESSION BG384217
 VERSION BG384217.1 GI:13308689
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 535)
 AUTHORS Fahrenkrug,S.C., Preking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
 and Keefe,J.W.
 TITLE Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 JOURNAL Unpublished (2000)
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCCAGTCACGACG
 Plate: 90 row: G column: 13
 Seq primer: ATTTAGTGACACTATAG.
 FEATURES source
 Location/Qualifiers
 1..535
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC LP1G"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."
 BASE COUNT 121 a 159 c 136 g 119 t

Query Match 10.9%; Score 39; DB 10; Length 535;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 220 ctgctggaagaaccttcttcagagaagaaca 258
 |||||||
 Db 311 CTGCTGGAAGAACCTCACTTTTCAGAGAAGACAAACA 349
 |||||||

RESULT 3
 AQ508395
 LOCUS 158 bp DNA linear GSS 29-APR-1999
 DEFINITION RPCI-11-274E11.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-274E11
 , DNA sequence.
 ACCESSION AQ508395
 VERSION AQ508395.1 GI:4713142
 KEYWORDS GSS.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 158)
 AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter
 ,J.C.
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other_GSSs: RPCI-11-274E11.TV
 Contact: Shaying Zhao, William Niernan, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genet cs (info@resgen.com). BAC end search page:
 http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.
 FEATURES Location/Qualifiers
 1..158
 /organism="Homo sapiens"
 /db_xref="GDB:7604938"
 /db_xref="taxon:9606"
 /clone="RPCI-11-274E11"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"
 BASE COUNT 58 a 12 c 69 g 19 t

Query Match 6.7%; Score 24; DB 12; Length 158;
 Best Local Similarity 100.0%; Pred. No. 0.37;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 326 ggaaggaggaggaggaggaggaggagg 349
 |||||||
 Db 120 GGAGGAGGAGGAGGAGGAGGAGGAG 143
 |||||||

RESULT 4
 Z44377
 LOCUS 292 bp mRNA linear EST 14-NOV-1994
 DEFINITION HSC1ZB081 normalized infant brain cDNA Homo sapiens cDNA clone
 c-1zb08, mRNA sequence.
 ACCESSION Z44377
 VERSION Z44377.1 GI:573506
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 292)
 AUTHORS Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
 ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
 Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
 Sebastiani-Kabaktchis,C. and Tessier,A.
 TITLE IMAGE: molecular integration of the analysis of the human genome
 and its expression
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 MEDLINE 95277534
 COMMENT Contact: Genethon
 Genexpress-Genethon

Genethon Centre de recherche sur le Genome Humain
1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpres@genethon.fr
Single read.
Genexpres library idt: C; Genexpres_sequence_idt: ylc-1zb08
Seq primer: (-21)M13 universal.
Location/Qualifiers

FEATURES

source

1..292
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-1zb08"
/clone_lib="normalized infant brain cdna"
/sex="female"
/tissue_type="total brain"
/dev_stages="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI; sex=female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 50 a 87 c 56 t 3 others

ORIGIN

Query Match 6.7%; Score 24; DB 10; Length 292;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 agctctggcgtgccttcacagg 154

|||||

DB 236 AGCTCTGGCGCTGCTCCAGG 259

RESULT 5
AQ528018/c

LOCUS AQ528018 709 bp DNA linear GSS 18-MAY-1999
DEFINITION RPCI-11-313F19.TV RPCI-11 Homo sapiens genomic clone RPCI-11-313F19
, DNA sequence.

ACCESSION AQ528018

VERSION AQ528018.1 GI:4840172

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 709)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building

JOURNAL Unpublished (1997)

COMMENT Other GSSs: RPCI-11-313F19.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

FEATURES

source

1..709
/organism="Homo sapiens"
/db_xref="GDB:761946"
/db_xref="taxon:9606"
/clone="RPCI-11-313F19"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 159 a 157 c 100 g 293 t
ORIGIN

Query Match 6.7%; Score 24; DB 12; Length 709;

Best Local Similarity 100.0%; Pred. No. 0.44; 0; Indels 0; Gaps 0;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggag 349

|||||

DB 84 GGAGGAGGAGGAGGAGGAGGAG 61

RESULT 6

AG163359/c

LOCUS AG163359 716 bp DNA linear GSS 09-JAN-2002

DEFINITION Pan troglodytes DNA, clone: RP43-030B04.T7, genomic survey

sequence.

ACCESSION AG163359

VERSION AG163359.1 GI:16693037

KEYWORDS GSS; GSS (genome survey sequence).

SOURCE Pan troglodytes male lymphocytes DNA, clone: RP43-030B04.T7.

Male BAC Library clone: RP43-030B04.T7.

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (sites)

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library RPCI-43

Unpublished

2 (bases 1 to 716)

REFERENCE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

TITLE Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimbges@gsc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>,

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library RPCI-43 This BAC

end was generated during the R&D process and may have higher chance

of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACE3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

Location/Qualifiers

1..716

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="RP43-030B04.T7"

/sex="male"

/cell_type="lymphocytes"

/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

BASE COUNT 117 a 246 c 84 g 268 t

ORIGIN

Query Match 6.7%; Score 24; DB 12; Length 716;

Best Local Similarity 100.0%; Pred. No. 0.44;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
|||||

Db 157 GGAGGAGGAGGAGGAGGAGGAAG 134

RESULT 7

AI394181
LOCUS tg67a04.xl Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:2113806
DEFINITION 3', mRNA sequence.

ACCESSION AI394181
VERSION AI394181.1 GI:4223728

KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 401)
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 480 Std Error: 0.00
Seq primer: -40UP from Gibco.

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 480 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES

source
1..401
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2113806"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"

/note="Organ: mixed (see below); Vector: pT73b-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2N6HM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 140 a 61 c 92 g 108 t
ORIGIN

Query Match 6.4%; Score 23; DB 9; Length 401;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 gagaggaggaggaggaggaggagg 349
|||||

Db 358 GAGGAGGAGGAGGAGGAGGAAG 380

RESULT 8

B20638
LOCUS B20638 827 bp DNA linear GSS 16-SEP-1997
DEFINITION F5016-T7 IGF Arabidopsis thaliana genomic clone F5016, DNA
sequence.

ACCESSION B20638
VERSION B20638.1 GI:2395692

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS

TITLE

JOURNAL

COMMENT

BAC End Sequences at ATGC

Unpublished (1997)

Other GSSs: F15016-Sp6

Contact: Ecker J.

Arabidopsis thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA

19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 175

High quality sequence stop: 226.

FEATURES

source

1..827

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="F15016"

/clone_lib="IGF"

/sex="hermaphrodite"

/note="Vector: BelOBA2II; Site_1: EcoRI; Site_2: EcoRI;

Produced by Thomas Altmann"

BASE COUNT 197 a 233 c 90 g 284 t

ORIGIN

Query Match 6.4%; Score 23; DB 12; Length 827;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 gagaggaggaggaggaggagg 349
|||||

Db 791 GAGGAGGAGGAGGAGGAAG 769

RESULT 9

BE050974/c

LOCUS

DEFINITION

za71c12.g51 Maize Glume cDNAs Library Zea mays cDNA clone za71c12

5', mRNA sequence.

ACCESSION BE050974

VERSION BE050974.1 GI:8368029

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

1 (bases 1 to 282)

O'Shaughnessy, A.L., Habermann, K., de la Bastide, M., Huang, E.N.,

Nascimento, L.O., Schutz, K., Matero, A., Swaby, I., See, L.-H., Preston

, R.R., Rodriguez, M.A., Shah, R.S., Shekher, M., Spiegel, L.A., Vil

, M.D., Dedhia, N.N. and McCombie, W.R.

Expressed sequence tags from Zea mays (maize)

Unpublished (2000)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: za71 row: c column: 12

Seq primer: -40M13RevUniv

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 827)

Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and

Ecker, J.

BAC End Sequences at ATGC

Unpublished (1997)

Other GSSs: F15016-Sp6

Contact: Ecker J.

Arabidopsis thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA

19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 175

High quality sequence stop: 226.

High quality sequence stop: 282.

FEATURES
source

Location/Qualifiers
1. .282
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="za71c12"
/clone_lib="Maize Glume cDNAs Library"

/note="Vector: lambda Zap II (Stratagene); Site_1: XhoI;
Site_2: EcoRI; Resistance: Ampicillin Autoexcision:
pBluescript SK (+/-) Titer: 7 x 10e-9 pfu/mL (as of
9/28/94)."

36 a 123 c 75 g 48 t

BASE COUNT
ORIGIN

Query Match 6.2%; Score 22; DB 9; Length 282;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 ggaggaggaggaggaggaggc 350
|||||
Db 50 GGAGGGAGGGAGGAGGAAGC 29

RESULT 10
BH293527

LOCUS CH230-30L8.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION BH293527 309 bp DNA linear GSS 30-NOV-2001
CH230-30L8, DNA sequence.

ACCESSION BH293527 GI:17205935
VERSION BH293527
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 309)
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn
A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de
Jong, P. and Fraser, C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other_GSS: CH230-30L8.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(<http://www.chori.org/bacpac/orering-information.htm>). BAC end

page: http://www.tigr.org/tcdr/bac_ends/rat/bac_end_intro.html

Plate: 30 row: L column: 8

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. .309

/organism="Rattus norvegicus"

/strain="BN/SsNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-30L8"

/clone_lib="CHORI-230 Segment 1"

/sex="Female"

/cell_type="Brain"

/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;

CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by

Pieter de Jong"

114 a 36 c 86 g 73 t

BASE COUNT

ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 aggaggaggaggaggaggagg 349
|||||
Db 278 AGGAGGGAGGGAGGAGGAAG 299

RESULT 11
BF592428/c

LOCUS BF592428 371 bp mRNA linear EST 12-DEC-2000
DEFINITION 7156a03.x1 NCI_CGAP_Br16 Homo sapiens cDNA clone IMAGE:338668 3'
similar to contains element TARI TARI repetitive element ;, mRNA
sequence.

ACCESSION BF592428 GI:11684752
VERSION BF592428
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 371)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Trieste Procurement: Dennis Sgroi, M.D., Kristina Cole, M.D., Ph.D.

student, Michael R. Emmert-Buck, M.D., Ph.D., Ph.D. Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -4ORF from Gibco.

Location/Qualifiers

1. .371

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:338668"

/clone_lib="NCI_CGAP_Br16"

/sex="female"

/tissue_type="lobullar carcinoma in situ"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: Breast; Vector: pAMP1; mRNA made from breast

carcinoma tissue, cDNA made by oligo-dT priming.

Directionally cloned. Size-selected on agarose gel,

average insert size 400 bp. Primary library,

non-amplified."

20 a 184 c 20 g 147 t

BASE COUNT
ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 aggaggaggaggaggaggagg 349
|||||
Db 166 AGGAGGGAGGGAGGAGGAAG 145

RESULT 12
R11136

LOCUS R11136 406 bp mRNA linear EST 11-APR-1995
DEFINITION Yf39b09.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone

IMAGE:129209 5' similar to gb:L21696_cds1 PROTHYMOSIN ALPHA (HUMAN)
); contains MER22 repetitive element i; mRNA sequence.

R11136
R11136.1 GI:763871

EST.
SOURCE
ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 406)
Hillier, L., Clark, N., Duboue, T., Elliston, K., Hawkins, M., Holman
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson R

TITLE
JOURNAL
COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 2665

High quality sequence stops: 290 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2665 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 290.

FEATURES

Location/Qualifiers

1..406

/organism="Homo sapiens"
/db_xref="GDB:481370"
/db_xref="taxon:9606"
/clone="IMAGE:129209"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGAATTAATAAGATCTTTTTTTTTTTTTTTT 3']/
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN

101 a 100 c 131 g 72 t
2 others

Query Match 6.2%; Score 22; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggagggaggaggaggaggagg 349

Db 356 AGGAGGGAGGGAGGAGGAGGAG 377

RESULT 13
AI481405/c

LOCUS
DEFINITION
AI481405 465 bp mRNA linear EST 09-MAR-1999
v917d11.x1 Soares mouse NBMH Mus musculus cDNA clone IMAGE:861621

ACCESSION
AI481405

VERSION
AI481405.1 GI:4374631

KEYWORDS
EST.

SOURCE
house mouse.

ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 465)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

TITLE
JOURNAL
COMMENT

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:505709
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
High quality sequence stop: 440.

FEATURES

source

1..465

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:861621"
/clone_lib="Soares mouse NBMH"
/sex="male"
/tissue_type="heart"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGGAGAAAGTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT 106 a 152 c 93 g 114 t

ORIGIN

Query Match 6.2%; Score 22; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggagggaggaggaggaggagg 349

Db 63 AGGAGGGAGGGAGGAGGAGGAG 42

RESULT 14
BE050973/c

LOCUS

DEFINITION
za71c12.b50 Maize Glume cDNAs Library Zea mays cDNA clone za71c12
5', mRNA sequence.

ACCESSION
BE050973

VERSION
BE050973.1 GI:8368028

KEYWORDS
EST.

SOURCE
Zea mays.

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 (bases 1 to 478)

AUTHORS

O'Shaughnessy, A.L., Habermann, K., de la Bastide, M., Huang, E.N.,
Nascimento, L.U., Schutz, K., Matero, A., Swaby, I., See, L.-H., Preston
R.R., Rodriguez, M.A., Shah, R.S., Shekher, M., Spiegel, L.A., Vil
M.D., Dedhia, N.N. and McCombie, W.R.
Expressed sequence tags from Zea mays (maize)

JOURNAL
COMMENT

Unpublished (2000)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org

Plate: za71 row: c column: 12
Seq primer: -40M3ForUniv
High quality sequence stop: 478.
Location/Qualifiers

FEATURES
source

1. .478
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="za71c12"
/clone_lib="Maize Glume cDNAs Library"
/note="Vector: Lambda zap II (Stratagene); Site_1: XhoI;
Site_2: EcoRI; Resistance: Ampicillin Autoexcision;
pBluescript SK (+/-) Titer: 7 x 10e-9 pfu/mL (as of
9/28/94)."

BASE COUNT 89 a 170 c 127 g 92 t
ORIGIN

Query Match 6.2%; Score 22; DB 9; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.1; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggc 350
|||||
DB 44 GGAGGGAGGGAGGAGGAGC 23

RESULT 15
AZ654841/c

LOCUS BH337839/c
DEFINITION BH337839 TV CHORI-230 Segment 1 Rattus norvegicus genomic clone UUGC1M0529P18 F, DNA sequence.

ACCESSION AZ654841
VERSION 1
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 483)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D. Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0529 row: P column: 18
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 483.
Location/Qualifiers

FEATURES
source

1. .483
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0529P18"

/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 107 a 140 c 79 g 157 t
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 483;
Best Local Similarity 100.0%; Pred. No. 4.1; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggc 347
|||||
DB 340 GGAGGGAGGGAGGAGGAGGA 319

RESULT 16
BH337839/c

LOCUS BH337839/c
DEFINITION BH337839 TV CHORI-230 Segment 1 Rattus norvegicus genomic clone CH230-194J9, DNA sequence.

ACCESSION BH337839
VERSION 1
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 515)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment Unpublished (1999)
Other_GSSs: CH230-194J9.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pjejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or ering_information.htm). BAC end plate: 194 row: J column: 9
Seq primer: T7
Class: BAC ends.
Location/Qualifiers

1. .515

FEATURES
source

```

/organism="Rattus norvegicus"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-194J9"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/notes="Vector: pTARAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT      152 a 154 c 90 g 119 t
ORIGIN

Query Match      6.2%; Score 22; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 gagagaggaggaggaggagga 347
|||||
DB 243 GGAGGAGGAGGAGGAGGA 222

RESULT 17
LOCUS BM310687/c 523 bp mRNA linear EST 03-JAN-2002
DEFINITION ig47b12.y1 HR85 islet Homo sapiens cDNA 5', mRNA sequence.
ACCESSION BM310687
VERSION BM310687.1 GI:18045003
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ig47b12.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -408P from Gibco
High quality sequence stop: 488.
FEATURES
source
1. 523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:

```

```

BASE COUNT      109 a 127 c 88 g 199 t
ORIGIN

Query Match      6.2%; Score 22; DB 10; Length 523;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 gagagaggaggaggaggagga 348
|||||
DB 291 GAGGAGGAGGAGGAGGA 270

RESULT 18
LOCUS AZ647100/c 525 bp DNA linear GSS 14-DEC-2000
DEFINITION 1M0513122F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0513122 F, DNA sequence.
ACCESSION AZ647100
VERSION AZ647100.1 GI:11778230
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 525)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0513 row: L column: 22
Seq primer: CTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 525.
FEATURES
source
1. 525
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0513122"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42hv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi14732114|gbAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into

```

chemically-competent *E. coli* XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

95 a 156 c 91 g 183 t

Query Match 6.2%; Score 22; DB 12; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 aggaggaggaggaggaggagg 349
|||||

Db 347 AGGAGGAGGAGGAGGAGGAG 326

RESULT 19
LOCUS

AQ415718 544 bp DNA linear GSS 23-MAR-1999
DEFINITION RPCI-11-204G24-TV RPCI-11 Homo sapiens genomic clone RPCI-11-204G24
, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

AQ415718.1 GI:4474687
GSS.

ORGANISM

human.

REFERENCE

AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building

JOURNAL

COMMENT

Other_GSSs: RPCI-11-204G24.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: kbe@igr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genet cs (inforesgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..544

/organism="Homo sapiens"

/db_xref="GDB:7578119"

/db_xref="taxon:9606"

/clone="RPCI-11-204G24"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

RPCI11 Human Male BAC Library"

90 a 162 c 78 g 214 t

BASE COUNT

ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 544;

Best Local Similarity 100.0%; Pred. No. 4.2;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 aggaggaggaggaggaggagg 349

|||||

Db 182 AGGAGGAGGAGGAGGAGGAG 161

RESULT 20
LOCUS

AV938157/c

DEFINITION

AV938157 K. Sato unpublished cDNA library, strain H602 adult,
heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bahl6g20 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

Hordeum vulgare subsp. spontaneum.

Hordeum vulgare subsp. spontaneum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Hordeum.

1 (bases 1 to 552)

Sato,K., Saisho,D. and Takeda,K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..552

/organism="Hordeum vulgare subsp. spontaneum"

/strain="H602"

/db_xref="taxon:77009"

/clone="bahl6g20"

/clone_lib="K. Sato unpublished cDNA library, strain H602

adult, heading stage top three leaves"

/tissue_type="top three leaves"

/dev_stage="adult, heading stage"

106 a 175 c 142 g 129 t

BASE COUNT

ORIGIN

Query Match 6.2%; Score 22; DB 9; Length 552;

Best Local Similarity 100.0%; Pred. No. 4.2;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggaggaggaggaggaggagg 347

|||||

Db 37 GGAGGAGGAGGAGGAGGAGGA 16

RESULT 21

LOCUS

AQ078910/c

DEFINITION

AQ078910

CIT-HSP-2367J11.TR CIT-HSP Homo sapiens genomic clone 2367J11, DNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 569)

Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,

Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and

Venter,J.C.

Use of a random human BAC End Sequence Database for Sequence-Ready

Map Building

Unpublished (1998)

Other_GSSs: CIT-HSP-2367J11.TF

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1. 569
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2367J11"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
 HindIII"

BASE COUNT 68 a 189 c 57 g 254 t 1 others
 ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 569;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggaggaggaggaggaggagg 349
 |||||
 Db 454 AGGAGGAGGAGGAGGAGGAGG 433

RESULT 22
 AZ378099/c 572 bp DNA linear GSS 02-OCT-2000
 LOCUS
 DEFINITION
 clone UUGCLM0132P08 R, DNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Mus musculus
 house mouse.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 572)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

TITLE
 JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0132 row: P column: 08
 Seq primer: CACACAGGAAACACTATGACC
 Class: plasmid ends
 High quality sequence stop: 572.

FEATURES
 source
 Location/Qualifiers
 1. 572
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0132P08"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (gil4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 130 a 137 c 118 g 187 t
 ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 572;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggaggaggaggaggaggagg 349
 |||||
 Db 231 AGGAGGAGGAGGAGGAGGAGG 210

RESULT 23
 AZ828464/c 591 bp DNA linear GSS 20-FEB-2001

LOCUS
 DEFINITION
 clone UUGC2M0105N23 F, DNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Mus musculus
 house mouse.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 591)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0103 row: N column: 23
 Seq primer: CGTGTAAACACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 591.

FEATURES
 source
 Location/Qualifiers
 1. 591
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0105N23"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pWD42nv: Purified genomic DNA from Mus musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 117 a 189 c 110 g 175 t
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 591;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaaggaggaggaggaggaggc 350
|||||
Db 130 GGAGGGAGGGAGGAGGAGGAGC 109

RESULT 24
AQ924975/c
LOCUS
DEFINITION AQ924975 615 bp DNA linear GSS 21-DEC-1999
RPCI-23-277E11-TV RPCI-23 Mus musculus genomic clone RPCI-23-277E11
, DNA sequence.
ACCESSION AQ924975.1 GI:6613978
VERSION
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 615)
AUTHORS Zhao,S., Nieman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 277 row: E column: 11
Seq primer: T7
Class: BAC ends.

FEATURES
source
1. .615
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-277E11"
/clone_lib="RPCI-23"

/sex="female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 144 a 158 c 101 g 212 t
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 615;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggaggaggaggaggaggaggc 349
|||||
Db 506 AGGAGGGAGGAGGAGGAGGAG 485

RESULT 25
B70930/c
LOCUS
DEFINITION B70930 623 bp DNA linear GSS 21-JUN-1998
CIT-HSP-2063G10.TF CIT-HSP Homo sapiens genomic clone 2063G10, DNA sequence.

ACCESSION B70930
VERSION B70930.1 GI:2710154
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 623)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-2063G10.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
1. .623
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:706153"
/db_xref="taxon:9606"
/clone="2063G10"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 74 a 209 c 72 g 268 t
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 623;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agggaggaggaggaggaggagg 349
 |||||||||||||||||||
 Db 491 AGGAGGAGGAGGAGGAGGAG 470

RESULT 26
 BF984535
 LOCUS 602307723F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4399276 5',
 DEFINITION mRNA sequence.
 ACCESSION BF984535
 VERSION BF984535.1 GI:12387347
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 673)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CNA Library Preparation: Life Technologies, Inc.
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLML at:
<http://image.llnl.gov>
 plate: LLAM10102 row: m column: 05
 High quality sequence stop: 671.

FEATURES
 source
 1..673
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4399276"
 /clone_lib="NIH_MGC_88"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: small intestine; Vector: pCMV-SPORT6;
 Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
 oligo-dT primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 150 a 184 c 204 g 135 t

Query Match 6.2%; Score 22; DB 10; Length 573;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 gggaggaggaggaggaggaggagg 347
 |||||||||||||||||||
 Db 33 GGAGGAGGAGGAGGAGGAGGA 54

RESULT 27
 AZ742461/c
 LOCUS AZ742461 RPCI-24-74G13.TJ RPCI-24 Mus musculus genomic clone RPCI-24-74G13,
 DEFINITION DNA sequence.
 ACCESSION AZ742461
 VERSION AZ742461.1 GI:12521331
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 713)
 AUTHORS Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B.,
 Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E.,

Russell, D., de Jong, P. and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-24
 Unpublished (1999)
 Other GSSs: RPCI-24-74G13.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@email.cho.org). Clones may be purchased from BACPAC
 resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
 page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 74 row: G column: 13
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..713
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-24-74G13"
 /clone_lib="RPCI-24"
 /sex="Male"
 /cell_type="Spleen/Brain"
 /note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;
 RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
 library was cloned in the pTARBAC1 cloning vector at the
 BamHI sites using MboI partially digested male C57BL/6J
 DNA."

BASE COUNT 126 a 225 c 114 g 248 t

Query Match 6.2%; Score 22; DB 12; Length 713;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agggaggaggaggaggaggagg 349
 |||||||||||||||||||
 Db 446 AGGAGGAGGAGGAGGAGGAAG 425

RESULT 28
 AQ983682/c
 LOCUS AQ983682 754 bp DNA linear GSS 30-JAN-2000
 DEFINITION RPCI-23-323J22.TJ RPCI-23 Mus musculus genomic clone RPCI-23-323J22
 , DNA sequence.
 ACCESSION AQ983682
 VERSION AQ983682.1 GI:6816887
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 754)
 AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret,
 B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.,
 and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from BACpac Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (inforesgen.com). BAC end page: http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html
 Plate: 323 row: j' column: 22
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..754
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-323J22"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: ECORI; Site_2: ECORI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of ECORI and ECORI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the ECORI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
 182 a 189 c 131 g 251 t 1 others

BASE COUNT

182 a 189 c 131 g 251 t 1 others

Query Match 6.2%; Score 22; DB 12; Length 754;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agagaggaggaggaggaggag 349
 Db 510 AGGAGGAGGAGGAGGAGGAG 489

RESULT 29

BE877424 760 bp mRNA linear EST 20-OCT-2000
 LOCUS 601485462F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888134 5',
 DEFINITION mRNA sequence.

ACCESSION BE877424
 VERSION BE877424.1 GI:10326200
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 760)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DNP/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9667 row: k column: 15

High quality sequence stop: 678.

FEATURES

Location/Qualifiers
 1..760
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3888134"
 /clone_lib="NIH_MGC_69"
 /tissue_type="large cell carcinoma, undifferentiated"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.1 kb. Library constructed by Life Technologies."

BASE COUNT 173 a 207 c 233 g 147 t
 ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 760;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggagaggaggaggaggaggagga 347
 Db 14 GGAGGAGGAGGAGGAGGAGGA 35

RESULT 30

BE617281 907 bp mRNA linear EST 20-OCT-2000
 LOCUS 601441977F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846235 5',
 DEFINITION mRNA sequence.

ACCESSION BE617281

VERSION BE617281.1 GI:9888219

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 907)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9558 row: i column: 20

High quality sequence start: 8

High quality sequence stop: 660.

FEATURES

Location/Qualifiers
 1..907
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3846235"
 /clone_lib="NIH_MGC_65"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

BASE COUNT 197 a 252 c 304 g 154 t
 ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 907;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggagaggaggaggaggaggagga 347
 Db 15 GGAGGAGGAGGAGGAGGAGGA 36

RESULT 31

BG823588 926 bp mRNA linear EST 22-MAY-2001
 LOCUS 602728977F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4868443 5',
 DEFINITION mRNA sequence.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM1233 row: f column: 24
High quality sequence stop: 668.

FEATURES
source
1. .1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4546583"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 261 a 294 c 332 g 181 t
ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 1068;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 326 ggaggaggaggaggaggaggga 347
|||||
Db 6 GGAGGAGGAGGAGGAGGA 27
|||||

RESULT 33
AZ995124/c
LOCUS
DEFINITION
AZ995124 43 bp DNA linear GSS 27-APR-2001
2M0280018R Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0280018 R, DNA sequence.

ACCESSION AZ995124
VERSION AZ995124.1 GI:13866351
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 43)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0280 row: 0 column: 18
Seq primer: CACACAGGAACAGCATGACC
Class: plasmid ends
High quality sequence stop: 43.
FEATURES
source
1. .43
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0280018"

```

/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      0 a      27 c      2 g      14 t
ORIGIN

```

Query Match 5.9%; Score 21; DB 12; Length 43;

Best Local Similarity 100.0%; Pred. No. 10; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349

Db 34 GGAGGAGGGGAGGAGGAG 14

RESULT 34

AZ351915

LOCUS

DEFINITION 1M0090K10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0090K10 F, DNA sequence.

ACCESSION AZ351915

VERSION AZ351915.1

KEYWORDS GI:10431152

SOURCE GSS.

ORGANISM house mouse.

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0090 row: K column: 10

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 51.

Location/Qualifiers

1..51

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

FEATURES

source

```

/clone_lib="UUGC1M0090K10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      16 a      0 c      35 g      0 t
ORIGIN

```

Query Match 5.9%; Score 21; DB 12; Length 51;

Best Local Similarity 100.0%; Pred. No. 10; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349

Db 19 GGAGGAGGGGAGGAGGAG 39

RESULT 35

AZ500262

LOCUS

DEFINITION 1M0338H04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0338H04 R, DNA sequence.

ACCESSION AZ500262

VERSION AZ500262.1

KEYWORDS GI:10679897

SOURCE GSS.

ORGANISM house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0338 row: H column: 04

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 62.

Location/Qualifiers

1..62

/organism="Mus musculus"

/strain="C57BL/6J"

FEATURES

source

/db_xref="taxon:10090"
 /clone="UUGC1M0338H04"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 27 a 1 c 32 g 2 t

BASE COUNT
 ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 62;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 ggaggaggaggaggaggaggagg 349
 |||||
 Db 26 GGAGGGAGGGAGGAGGAAG 46

RESULT 36
 AZ411857/c
 LOCUS
 DEFINITION 66 bp DNA linear GSS 03-OCT-2000
 Clone UUGC1M0185006 F, DNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 66)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE
 JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0185 Row: 0 Column: 06
 Seq primer: CGTTGTAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 66.
 Location/Qualifiers
 1. .66
 /organism="Mus musculus"

/strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0185006"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 3 a 37 c 2 g 24 t

BASE COUNT
 ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 66;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 ggaggaggaggaggaggaggagg 349
 |||||
 Db 57 GGAGGGAGGGAGGAGGAAG 37

RESULT 37
 AZ838528/c
 LOCUS
 DEFINITION 90 bp DNA linear GSS 20-FEB-2001
 Clone UUGC2M0134E17 F, DNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 90)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITLE
 JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0134 Row: E Column: 17
 Seq primer: CGTTGTAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 90.
 Location/Qualifiers
 1. .90
 /source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0134E17"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      7 a      49 c      0 g      34 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 90;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349
|||||
Db 42 GGAGGGAGGGAGGGAAGGAG 22

RESULT 38
AZ743191
LOCUS
DEFINITION      120 bp DNA linear GSS 25-JAN-2001
PC1-24-146C22.TJ RPCI-24 Mus musculus genomic clone RPCI-24-146C22
, DNA sequence.
ACCESSION      AZ743191
VERSION        AZ743191.1 GI:12522805
KEYWORDS       GSS.
SOURCE         house mouse
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 120)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other_GSSs: RPCI-24-146C22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/db/bac\_ends/mouse/bac\_end\_intro.html
Plate: 146 row: C column: 22
Seq primer: SP6
Class: BAC ends.

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0134E17"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      7 a      49 c      0 g      34 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 120;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349
|||||
Db 84 GGAGGGAGGGAGGGAAGGAG 104

RESULT 39
AI053852/c
LOCUS
DEFINITION      145 bp mRNA linear EST 15-JUL-1998
qi70c08.x1 NCI-CGAP_Ov26 Homo sapiens cDNA clone IMAGE:1861838 3'
similar to contains element TAR1 TAR1 repetitive element ;, mRNA
sequence.
ACCESSION      AI053852
VERSION        AI053852.1 GI:3321639
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Monica Brown, M.D., Elise Kohn, M.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers
1. .145
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1861838"
/clone_lib="NCI-CGAP_Ov26"
/sex="female"
/tissue_type="papillary serous carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pAMPl; mRNA made from
papillary serous ovarian carcinoma, cDNA made by oligo-dT
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 500 bp. Primary library,
non-amplified."
BASE COUNT      1 a      57 c      4 g      83 t
ORIGIN

```

```

Query Match      5.9%; Score 21; DB 9; Length 145;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 ggaggaggaggaggaggagg 349
|||||
Db 89 GGAGGGAGGGAGGAGGAAG 69

RESULT 40
AZ091117
LOCUS      161 bp      DNA      linear      GSS 08-MAY-2000
DEFINITION RPCI-23-7H17.TJ RPCI-23 Mus musculus genomic clone RPCI-23-7H17,
DNA sequence.
ACCESSION  AZ091117
VERSION     AZ091117.1 GI:7733160
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 161)
AUTHORS    Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
TITLE      Mouse BAC End Sequences from Library RPCI-23
JOURNAL
COMMENT    Other_GSSs: RPCI-23-7H17.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 7 row: H column: 17
            Seq primer: SP6
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..161
     /organism="Mus musculus"
     /strain="C57BL/6J"
     /db_xref="taxon:10090"
     /clone="RPCI-23-7H17"
     /clone_lib="RPCI-23"
     /sex="Female"
     /lab_host="DH10B"
     /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
     brain genomic DNA was isolated and partially digested
     with a combination of EcoRI and EcoRI Methylase. Size
     selected DNA was cloned into the pBAC3.6 vector at the
     EcoRI sites. The ligation products were transformed into
     DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT  42 a 16 c 45 g 58 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 161;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 gagaggaggaggaggagg 347
|||||
Db 134 GAGAGGGAGGAGGAGGAAG 154

Query Match      5.9%; Score 21; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 gagaggaggaggaggagg 347
|||||

```

```

RESULT 41
AZ457379
LOCUS      191 bp      DNA      linear      GSS 04-OCT-2000
DEFINITION 1M0260L10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0260L10 R, DNA sequence.
ACCESSION  AZ457379
VERSION     AZ457379.1 GI:10615504
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 191)
AUTHORS    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0260 row: L column: 10
            Seq primer: CACACAGGAGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 191.

FEATURES             Location/Qualifiers
     source           1..191
     /organism="Mus musculus"
     /strain="C57BL/6J"
     /db_xref="taxon:10090"
     /clone="UUGC1M0260L10"
     /clone_lib="Mouse 10kb plasmid UUGC1M library"
     /sex="Male"
     /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
     /note="Vector: PWD42nv; Purified genomic DNA from M.
            musculus C57BL/6J (male) was obtained from the Jackson
            Laboratory Mouse DNA Resource
            (http://www.jax.org/resources/documents/dnares/). The DNA
            was hydrodynamically sheared by repeated passage through a
            0.005 inch orifice at constant velocity. The sheared DNA
            was blunt end-repaired with T4 DNA polymerase and T4
            polynucleotide kinase. Adaptor oligonucleotides were
            ligated to the blunt ends in high molar excess. The
            adapted DNA was purified and size-selected for a 9.5 to
            10.5 kb range using preparative agarose gel
            electrophoresis. Vector DNA was prepared from a derivative
            of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
            inducible derivative of plasmid R1. The vector was ligated
            with adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            adapted vector DNA, and transformed into
            chemically-competent E. coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
BASE COUNT  46 a 55 c 44 g 46 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 gagaggaggaggaggagg 347
|||||

```


Db 169 GAGGAGGAGGAGGAGGA 189

RESULT 42

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ656509 212 bp DNA linear GSS 14-DEC-2000
1M0532002F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0532002 F, DNA sequence.

ACCESSION AZ656509
VERSION
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 212)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0532 row: 0 column: 02
Seq primer: CGTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 212.

Location/Qualifiers
1. .212
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0532002"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gii4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 39 a 83 c 43 g 47 t

ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 212;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349

Db 169 GAGGAGGAGGAGGAGGA 189

RESULT 43

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 269)
Aufrey, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes,
M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lam, Y.B., Lorenzo, F.,
Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastiani-Kabaktchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read. 19T removed at sequence 5' end
Genexpress.library.idt: C: Genexpress_sequence_idt: a3c-2te03
Seq primer: (-21)M13-universal.

Location/Qualifiers
1. .269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2te03"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site: 1: HindIII;
Site 2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 93 a 34 c 107 g 32 t

ORIGIN

Query Match 5.9%; Score 21; DB 10; Length 269;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349

Db 207 GGAGGAGGAGGAGGAGGAAG 227

RESULT 44

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 294)
Azaiz, J., Bouchier, C., Bois, F., Bouchier, C., da Silva, C., Devignes,
M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lam, Y.B., Lorenzo, F.,
Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastiani-Kabaktchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read. 19T removed at sequence 5' end
Genexpress.library.idt: C: Genexpress_sequence_idt: a3c-2te03
Seq primer: (-21)M13-universal.

Location/Qualifiers
1. .269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2te03"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site: 1: HindIII;
Site 2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 93 a 34 c 107 g 32 t

ORIGIN

Query Match 5.9%; Score 21; DB 10; Length 269;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349

Db 39 GGAGGAGGAGGAGGAGGAAG 19

RESULT 43

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 269)
Aufrey, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes,
M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lam, Y.B., Lorenzo, F.,
Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastiani-Kabaktchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read. 19T removed at sequence 5' end
Genexpress.library.idt: C: Genexpress_sequence_idt: a3c-2te03
Seq primer: (-21)M13-universal.

Location/Qualifiers
1. .269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2te03"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site: 1: HindIII;
Site 2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 93 a 34 c 107 g 32 t

ORIGIN

Query Match 5.9%; Score 21; DB 10; Length 269;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349

Db 207 GGAGGAGGAGGAGGAGGAAG 227

RESULT 44

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 294)
Azaiz, J., Bouchier, C., Bois, F., Bouchier, C., da Silva, C., Devignes,
M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lam, Y.B., Lorenzo, F.,
Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastiani-Kabaktchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read. 19T removed at sequence 5' end
Genexpress.library.idt: C: Genexpress_sequence_idt: a3c-2te03
Seq primer: (-21)M13-universal.

Location/Qualifiers
1. .269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2te03"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site: 1: HindIII;
Site 2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 93 a 34 c 107 g 32 t

ORIGIN

Query Match 5.9%; Score 21; DB 10; Length 269;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349

Db 207 GGAGGAGGAGGAGGAGGAAG 227

RESULT 44

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 294)
Azaiz, J., Bouchier, C., Bois, F., Bouchier, C., da Silva, C., Devignes,
M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lam, Y.B., Lorenzo, F.,
Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastiani-Kabaktchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read. 19T removed at sequence 5' end
Genexpress.library.idt: C: Genexpress_sequence_idt: a3c-2te03
Seq primer: (-21)M13-universal.

Location/Qualifiers
1. .269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2te03"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site: 1: HindIII;
Site 2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 93 a 34 c 107 g 32 t

ORIGIN

Query Match 5.9%; Score 21; DB 10; Length 269;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349

Db 207 GGAGGAGGAGGAGGAGGAAG 227

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 294)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112 USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0360 row: B column: 02
Seq primer: CGTTGTAACAGCGGCCAGT
Class: plasmid ends
High quality sequence stop: 294.
FEATURES
source
1. .294
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0360B02"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
/note="Vector: PRD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWB42 (gil14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 131 a 38 c 79 g 46 t
ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 294;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349
|||||
Db 188 GGAGGGAGGGAGGGAGGAAG 208

RESULT 45
B50482/c
LOCUS B50482 319 bp DNA linear GSS 20-JUN-1998
DEFINITION CIT-HSP-437M12.TP CIT-HSP Homo sapiens genomic clone 437M12, DNA
sequence.
ACCESSION B50482
VERSION B50482.1 GI:2602719

KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 319)
AUTHORS Kim, U.-J., Adams, M.D. and Simon, M.I.
TITLE Determination of clone end sequences of human Bacterial Artificial
Chromosomes
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-437M12.TV
Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source
1. .319
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:5398850"
/db_xref="taxon:9606"
/clone="437M12"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 47 a 119 c 21 g 132 t
ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 319;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349
|||||
Db 218 GGAGGGAGGGAGGGAGGAAG 198

Search completed: September 20, 2002, 04:07:24
Job time: 13778 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 20, 2002, 06:30:00 ; Search time 5250.46 Seconds
(without alignments)
880.830 Million cell updates/sec

Title: US-09-846-456-4
Perfect score: 221
Sequence: 1 gtaattgcgagcgagagtga.....aacacaaagtgtgaaacag 221

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sv.*
13: gb_un.*
14: gb_vl.*
15: em_be.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgtGO_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES
Result No. Score Match Length DB ID Description

1	221	100.0	221	6	AX351032	AX351032 Sequence
2	221	100.0	697	9	AF258627	AF258627 Homo sapi
3	221	100.0	1167	9	HA252201	HA252201 Homo sapi
4	221	100.0	1167	9	AF258623	AF258623 Homo sapi
5	221	100.0	3231	6	AX351029	AX351029 Sequence
6	221	100.0	7260	6	AX253452	AX253452 Sequence
7	221	100.0	96717	9	AL359182	AL359182 Human DNA
8	221	100.0	149034	9	AF275948	AF275948 Homo sapi
9	221	100.0	175064	2	AC012230	AC012230 Homo sapi
10	221	100.0	201144	9	AF287262	AF287262 Homo sapi
11	219	99.1	1750	9	AK022254	AK022254 Homo sapi
12	219	98.2	1556	9	AK024328	AK024328 Homo sapi
13	205	92.8	9854	6	AX127831	AX127831 Sequence
14	205	92.8	9854	6	AX139818	AX139818 Sequence
15	201	91.0	69570	2	AC021246	AC021246 Homo sapi
16	201	91.0	183999	6	AX092589	AX092589 Sequence
17	197	89.1	10442	6	AX060713	AX060713 Sequence
18	197	89.1	10442	6	AX060892	AX060892 Sequence
19	197	89.1	10442	9	AF285167	AF285167 Homo sapi
20	196	88.7	90698	2	AC021345	AC021345 Homo sapi
21	188	85.1	10474	6	AX060719	AX060719 Sequence
22	188	85.1	10474	6	AX060721	AX060721 Sequence
23	188	85.1	10474	6	AX060898	AX060898 Sequence
24	188	85.1	10474	6	AX060900	AX060900 Sequence
25	92	41.6	446	6	AX127764	AX127764 Sequence
26	92	41.6	446	6	AX139751	AX139751 Sequence
27	92	41.6	9741	6	AX127830	AX127830 Sequence
28	92	41.6	9741	6	AX139817	AX139817 Sequence
29	92	41.6	9741	6	AX351038	AX351038 Sequence
30	91	41.2	1643	6	AX060715	AX060715 Sequence
31	91	41.2	1643	6	AX060894	AX060894 Sequence
32	47	21.3	69570	2	AC021246	AC021246 Homo sapi
33	20	9.0	92227	9	AC097517	AC097517 Homo sapi
34	20	9.0	149710	2	AC022563	AC022563 Homo sapi
35	20	9.0	162496	2	AC027453	AC027453 Homo sapi
36	20	9.0	201938	2	AC012325	AC012325 Homo sapi
37	20	9.0	228703	2	AC097253	AC097253 Rattus no
38	19	8.6	735	6	A39868	A39868 Sequence 2
39	19	8.6	1641	9	AF056979	AF056979 Homo sapi
40	19	8.6	1773	14	SIVAGMGAGD	L19253 Simian immu
41	19	8.6	2064	9	HUMIFNRG	J03143 Human inter
42	19	8.6	2127	9	BC005333	BC005333 Homo sapi
43	19	8.6	2184	6	A30438	A30438 H.sapiens p
44	19	8.6	2715	7	STH022	X51962 Collipage H
45	19	8.6	4070	9	HSINFGAL	U19241 Homo sapien

ALIGNMENTS

RESULT	1	AX351032	Sequence 4 from Patent WO0183746.	221 bp	DNA	Linear	PAT 06-FEB-2002
LOCUS	AX351032	Sequence 4 from Patent WO0183746.					
DEFINITION	AX351032	Sequence 4 from Patent WO0183746.					
ACCESSION	AX351032	Sequence 4 from Patent WO0183746.					
VERSION	AX351032.1	GI:18616388					
KEYWORDS	human.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
TITLE	1 (sites)						
JOURNAL	Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,						
FEATURES	Regulatory nucleic acid sequences of the abcl gene						
source	Patent: WO 0183746-A 4 08-NOV-2001;						
	Avantis Pharma S.A. (FR)						
	Location/Qualifiers						
	1..221						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
BASE COUNT	44 a	62 c	73 g	42 t			
ORIGIN							

QY 61 ccgggctgcggcagggcaggggagctccgcgacacacagagccgggttcacggc 120
 |||||
 Db 987 CCGGGCTGGCAGGAGCGGGGGAGCTCCGGCACCACAGAGCCGGTTCACAGGC 1046
 |||||
 QY 121 gcttgctcctgtttttcccggttctgttttcccttcccgaaagcttgcac 180
 |||||
 Db 1047 GCTTTGCTCCTTGTGTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTC 1106
 |||||
 QY 181 ggggtaggagaaagagcgcacacacacaaagtggaaaacag 221
 |||||
 Db 1107 GGGGTAGGAGAAAGAGACGCAACAAAGTGGAAACAG 1147
 |||||

RESULT 4
 AF258623S1
 LOCUS 1167 bp DNA linear PRI 23-JUN-2000
 DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,
 promoter and exon 1.
 ACCESSION AF258623
 VERSION AF258623.2 GI:8677405
 KEYWORDS
 SEGMENT
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1167)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
 TITLE Analysis of hABC1 gene 5' end: additional peptide sequence,
 promoter region, and four polymorphisms
 JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
 REFERENCE 2 (bases 224 to 1167)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
 TITLE Direct Submission
 JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 REFERENCE 3 (bases 1 to 1167)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 REMARK
 COMMENT Sequence update by submitter
 FEATURES On Jun 23, 2000 this sequence version replaced gi:7769713.
 Source Location/Qualifiers
 1. 1167
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q31"
 224..844
 /gene="ABCA1"
 845..1147
 /number=1

promoter
 exon

BASE COUNT 278 a 313 c 328 g 248 t
 ORIGIN

Query Match 100.0%; Score 221; DB 9; Length 1167;
 Best Local Similarity 100.0%; Pred. No. 5.2e-118;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcagcagagtagtgaggccgggacccagcagagccgacacctctctc 60
 |||||
 Db 927 GTATTGGAGCGAGAGTAGTGGCGCGGACCCGACGAGCCGACCCCTTCTCTC 986
 |||||

QY 61 ccgggctgcggcagggcaggggagctccgcgacacacagagccgggttcacggc 120
 |||||

Db 987 CCGGGCTGGCAGGAGCGGGGGAGCTCCGGCACCACAGAGCCGGTTCACAGGC 1046
 |||||
 QY 121 gcttgctcctgtttttcccggttctgttttcccttcccgaaagcttgcac 180
 |||||
 Db 1047 GCTTTGCTCCTTGTGTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTC 1106
 |||||
 QY 181 ggggtaggagaaagagcgcacacacacaaagtggaaaacag 221
 |||||
 Db 1107 GGGGTAGGAGAAAGAGACGCAACAAAGTGGAAACAG 1147
 |||||

RESULT 5
 AX351029
 LOCUS 3231 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 1 from Patent WO0183746.
 ACCESSION AX351029
 VERSION AX351029.1 GI:18616385
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P.,
 Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
 TITLE Regulatory nucleic acid sequences of the abcl gene
 JOURNAL Patent: WO 0183746-A 1 08-NOV-2001;
 Aventis Pharma S.A. (FR)
 FEATURES Location/Qualifiers
 source 1. 3231
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 809 a 773 c 876 g 773 t
 ORIGIN

Query Match 100.0%; Score 221; DB 6; Length 3231;
 Best Local Similarity 100.0%; Pred. No. 5.3e-118;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcagcagagtagtgaggccgggacccagcagagccgacacctctctc 60
 |||||
 Db 2894 GTAATTGGAGCGAGTAGTGGCGCGGACCCGACGAGCCGACCCCTTCTCTC 2953
 |||||
 QY 61 ccgggctgcgagcagggcagggcagctccgcgacacacagagccgggttcacggc 120
 |||||
 Db 2954 CCGGGCTCGGAGCGGCGGAGGAGCTCCGGCACCACAGAGCCGGTTCACAGGC 3013
 |||||
 QY 121 gcttgctcctgtttttcccggttctgttttcccttcccgaaagcttgcac 180
 |||||
 Db 3014 GCTTTGCTCCTTGTGTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTTGTC 3073
 |||||
 QY 181 ggggtaggagaaagagcgcacacacacaaagtggaaaacag 221
 |||||
 Db 3074 GGGGTAGGAGAAAGAGACGCAACAAAGTGGAAACAG 3114
 |||||

RESULT 6
 AX253452
 LOCUS 7260 bp DNA linear PAT 10-OCT-2001
 DEFINITION Sequence 3 from Patent WO0170810.
 ACCESSION AX253452
 VERSION AX253452.1 GI:16073979
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 7260)
 AUTHORS Schmitz,G. and Bodzioch,M.
 TITLE Atp binding cassette transporter 1 (abcl) gene polymorphisms and
 uses thereof for the diagnosis and treatment of lipid,
 cardiovascular or inflammatory disorders

Jr.
Complete genomic sequence of the human ABCA1 gene: analysis of the
human and mouse ATP-binding cassette A promoter
Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
20345099
2 (bases 1 to 149034)
Santamarina-Fojo, S., Peterson, K.M., Knapper, C.L., Freeman, L.A.,
Remaley, A.T., Yang, X.-P., Haudenschild, C.C., Blackmon, E.E.,
Francois, T.L. and Brewer, H.B. Jr.
Direct Submission
Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA
FEATURES
Location/Qualifiers
1..149034
/organism="Homo sapiens"
/db_xref="taxon:9606"
Join(1454..1674,25831..25989,40385..40478,45012..45153,
46423..46541,67729..67850,70831..71007,83960..84052,
89010..89250,91962..92101,92433..92549,96758..96955,
97702..97907,98428..98604,100391..100613,102360..102581,
103642..105064,104951..105064,106862..107033,
108023..108154,109460..109602,109806..109943,
110646..110866,112124..112196,113183..113385,
115053..115101,115297..115410,116806..116954,
118604..118728,119964..120062,123094..123283,
124804..124898,126208..126240,127363..127468,
128943..129017,129539..129708,130939..131116,
133114..133229,133341..133485,134526..134649,
135736..135865,136130..136250,137038..137100,
138008..138114,140469..140610,140982..141116,
142060..142163,142646..142738,143397..143640,
144581..148034)
/gene="ABCA1"
5'UTR
Join(1454..1674,25831..25923)
/product="ABCA1"
gene
1454..148034
/gene="ABCA1"
repeat_region
2406..2678
/rpt_family="Alu"
repeat_region
4240..4509
/rpt_family="Alu"
repeat_region
4998..5270
/rpt_family="Alu"
repeat_region
5363..5839
/rpt_family="Alu"
repeat_region
6615..6877
/rpt_family="Alu"
repeat_region
8800..9046
/rpt_family="Alu"
repeat_region
9998..10279
/rpt_family="Alu"
repeat_region
11859..12133
/rpt_family="Alu"
repeat_region
12810..12902
/rpt_family="Alu"
repeat_region
13220..13399
/rpt_family="Alu"
repeat_region
15508..15787
/rpt_family="Alu"
repeat_region
18600..18860
/rpt_family="Alu"
repeat_region
20950..21206
/rpt_family="Alu"
repeat_region
25038..25314
/rpt_family="Alu"
CDS
Join(25944..25989,40385..40478,45012..45153,46423..46541,
67729..67850,70831..71007,83960..84052,89010..89250,
91962..92101,92433..92549,96758..96955,97702..97907,
98428..98604,100391..100613,102360..102581,103642..103846,
104951..105064,106862..107033,108023..108154,
109460..109602,109806..109943,110646..110866,
112124..112196,113183..113385,115053..115101,115297..115410,116806..116954,
118604..118728,119964..120062,123094..123283,
124804..124898,126208..126240,127363..127468,
128943..129017,129539..129708,130939..131116,
133114..133229,133341..133485,134526..134649,
135736..135865,136130..136250,137038..137100,
138008..138114,140469..140610,140982..141116,
142060..142163,142646..142738,143397..143640,
144581..148034)
/gene="ABCA1"
112124..112196,113183..113385,115053..115101,
115297..115410,116806..116954,118604..118728,
119964..120062,123094..123283,124804..124898,
126208..126240,127363..127468,128943..129017,
129539..129708,130939..131116,133114..133229,
133341..133485,134526..134649,135736..135865,
136130..136250,137038..137100,138008..138114,
140469..140610,140982..141116,142060..142163,
142646..142738,143397..143640,144581..144721)
/gene="ABCA1"
codon_start=1
/product="ABCA1"
protein_id="AAF86276.1"
db_xref="GI:9247086"
translation="MACPQLRLLLKLNLFRRRTQCLLELVANWPLFLLILISVRL
SYPIEHECHFPNKPASGLTFWVGIIICNANPCFRPTPGEAAGVGNFKSLV
ARLSDARRLLYSQKDTSMKMRKVLRLTQQIKSSNKLQDLVDNTEFSGFLYH
NLSPKSTVDKMLRADYLLHKVFLQGLHLTSLCNSKSEMIQLQDEVSFCLGLP
REKAAERVLRLSNMDILKPLRLNSTSPPKELAEATKTLHSLGTLAQELFSMR
SWSDMROEVMLTNVSSSTOIQAVSRIVCGHPHGGGLKIKSLAWEDNNYKALP
GGNGTEAETFPDNTTPYCNLNMKNLSSPLSRIWKALKPLLVKILLYTDPAT
RQVNAENKTFQELAVFHLGEMWELSPKIWFENSNQEMDLVRLMDSRDHFWFE
QQLDLQMTAQDIYAFALAKPEDVQSSNGSYTWREAFNETNQAIRTSRMEDCNLN
KLEPIATEVWLINKSMELLDERKFWAGIVFTGTPGIELPHVHYKIRMDINVERT
NKIKDGYMDPGPRADPFEDMRKVFYQDVQEQAIIRVLTKETKTKYVMQMPY
PCYVDDIFLRVMSRMLFTLWISYAVIIGIVYEKARLKEKTRIMGLDNLILW
PSWFISSILPLVNSAGLLVILKGNLLPYSDPSVVFVFLSVFVILQCFLLSTLE
SRANLAACGGIITYLPIPLVCLVADQYVFTLKIFAXLLSPVAFVPGGCEYFALFE
EQGIGQWMDNLFESPEEDGNLTYSMMFLDFTLYGVMWYEAVPFGQYGIPLPW
YFPTKSYWFGESDEKSHGPNKRISEICMEBEPTHLKIGVQNLVRYRDMKV
AVDGLNLFYEGQITSLFNGHAGKTTMSILTLGTPPTSGTAYILGKDIRSEMSIR
ONLGVCPQHNVLFDMLTVEHHIWFYARLKGISEKHAEMQMALDGLPSSKSKST
SQLSGGQRKLSVALAFVGSKVILDEPTAGVDPYSRGIWELLYKRGRTIILT
HMDREADVIGRIALILSHGLKCCVSSSLFLKQLGTGYTLVLYKKDVSSSCRNSS
STVYLKEDSVQSSSDGSHDESITLIDVSAISNLIRKHVSSEARLVEDIGHEL
TYVLPYAAKEGAFELFHEIDRLSDIGISSYGETTLEIFLVAEESGVDAETS
DGTLPARRNRRAFGDKQCLRPFTEDDAADPNDSIDIPESRETLLSGMDGKGYQVK
GWKLTQOQFVALLKRLLIARRSRKGFQIIVPAVFCIALVFLVPPGKYPSLE
LQPMYNEQVTFVSDAPEDPTGLLELNLTKDPGFCRCMEGNPIPTPCOAGEEHW
TAPVPTQIMDLFONGNWTMONPSPACQSSDKIKKMLPVCPPGAGGLPPORQNTA
DILQDLTGRNISLYLVITYVQIIAASLKNKIWNVEFRYGGFSLGVSNQALPPSQEVN
DAXQMKKHLKALDSSADRFNLGSLRMTGLDTRNNVYKMFNKNHAISSFLNLIN
NAILRANLQGENPSHYGIFATFNHPLNLTKOOLSEVAXMTTSDVLVSIYFAMSFV
PASVFLIQERSVSKAKHLOFISGVKPVYWLSEVMDMCNVVVPATVLIITFCFOQ
KSYVSSNLPVIALLLLYGHSITPLMYPASVFEKIPSTAYVLTSLNLTGINGSVA
TFVLELFTDNNLNNILKSLVFLIFPHFCGLRGLDLMVKNQAMADALERGFENVS
PLSWDLVRNLFFAMAVEGVWFFLTIVLIQYRFFIRPRVNAKLSPLNDEDEDVRRHQ
RLDGGQNDILEIKELTKIYRRKRKPAVDRCVIGIPGECFGLGVNGAGKSTFRM
LTGDTVTTRGDAFLNNSILSNIHEVHONMGYCPQFDAITELLTGREHVEFALLRGV
PEKVGKGEWAIRKGLGVKGEKAGYAGNYSGNKRKLTAMALIGGPPVFLDPTTC
MDPKARFLNICALSVYKGRSVLTSHSMECEALCTRMALMIVNGRFLGSLGVQHLK
NRFGDGTIVVRIAGSNPDLPQVDFGLAPPVGSXKXKKNMLIYQLPSSLSLARI
FSILSQSKKRLHIEDYSVSQTLQDVFVNFARQDSDDDDLKDLSLHKNQTVADVAVLT
SFLQDEKVKESYV"
29916..30180
/rpt_family="Alu"
31050..31314
/rpt_family="Alu"
31692..31973
/rpt_family="Alu"
36167..36440
/rpt_family="Alu"
37617..37892
/rpt_family="Alu"
37999..38166
/rpt_family="Alu"
38172..38269
/rpt_family="Alu"
38995..39271
/rpt_family="Alu"
41547..41565
/gene="ABCA1"
/note="A009X28"

VERSION AF287262.1 GI:13876612
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 211144)
 AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
 TITLE Human and mouse abca1 comparative sequencing and transgenesis studies revealing novel regulatory sequences
 JOURNAL Genomics 73 (1), 66-76 (2001)
 MEDLINE 21251004
 REFERENCE 2 (bases 1 to 201144)
 AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Rubin,E. and Cheng,J.F.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUL-2000) Genome Science Department, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley, CA 94720, USA
 FEATURES
 source
 1..201144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 join(33931..34151,58319..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,123239..123569,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150438,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170329,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..181457)
 /gene="ABCA1"
 /product="ATP-binding cassette 1 sub-family A member 1"
 33931..181457
 /gene="ABCA1"
 33931..34151
 /gene="ABCA1"
 /number=1
 58319..58476
 /gene="ABCA1"
 /number=2
 join(58411..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,123239..123569,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150438,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170329,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..178700)
 /gene="ABCA1"
 /codon_start=1
 /product="ATP-binding cassette 1 sub-family A member 1"
 /protein_id="AAK43526.1"
 /db_xref="GI:13876613"
 /translation="MACWQQLLLKLNLTFRRTQCQLLLEAVPLFIILISVRL
 SYPPEQHCHEPKNKPSAGLTPWQGLICNANNPCFRYPPTGAPGVGNFNKSIY
 ARFSDARLLYSQKDTSMKDMRWLRTLQIKKSSNNKLQDFVDFNFTSGFLYH
 NLSLPKSTVDKMLRADVILHKVFLQGYQLTSLCNGSKSEMIQLGQVEYSELGLP

KEKLAARVLRNMDILKILRTLNSTSPFSPSKELAEATKTLHLSIGTSLAQELFSMR
 SWSMDRQEVMLTNVNSSSTQIYQAVSRIVCGHPEGGGLAIKSLNWNEDNNYRALF
 GNGTDEDAETFDNSTTPYCNLMANLESSPLRIIWRKALPLLVKLLYTPDPTAT
 RQYMAEVNKTTFQELAVFLHLEGMEELSPKINTFMENSDMLRVKLLSRNDHFWE
 QDQDGLDWAQIDAVFLAKHEDVQSSSVTWTREAFNETQAKITISRFMECNLN
 KLEPIATEVWLINKSMELDERKFWAGIVFTGTPGSIELPHHVKYKIRMDINDVERT
 NKTGDYWPGRADPPEDMRYVWGFAVLDQVVEQAIIRVLTGTEKTKVYVMOOMPY
 PCYVDIFLRVMSRSPLEMTLAWIYSAVILKGIYVEKARLKEFTMRIMGDLNLSIF
 FSWFISSLPIILVAGLLVILKGLNLLPYSDPSVVFSLFVAVVTIIQCLLSLTLF
 SRANLACGGIIIFTLIPLVLCVAMQDYGFTLKIFASLLSPVAFGEGCBGFALFE
 EQGIVQWNLNLFESVEEDGNLTTSVSMLENTLYGYMTWYIEAVFGQYGIPIRPW
 YPCTKSYWFEESDEKSHPSNQRKRISEICMEEPHKLGLVSIONLVKYVRDGMKV
 AVDLALNFYEQITSLFNGNGAKTTTMSILTLGLTPPTSGTAYILGKDIRSEMSTIR
 QNLGVCPOHNVLFDMLTVEHIWFAVLKGLSKHKVKAEMEQALDVLGPPSKLSKT
 SOLSGMQKRLSVLAFVGSKVVIIDEPAGVDVPSRRIWELLILKYRQGTIILST
 HMDEADVLGDRIALIISHGKLCVGSSEFLKQNGTGYTLTKKDVESLSLSCNRS
 STVSYLKEESVQSSSDAGSHEDTLTIDVSAISNLIRKHVSEARLVEDIGHEL
 TIVLYPEAKEGAFELFHEIDRDLGSISSYSETTLEEFILKVAEESGDAETS
 DGTLPARRNRRAFGDKSCLRPFTEDDAADPNDSIDIPESRETDLISGMDGKGSQVK
 GWKLTQOQFVALLIKRLLIARRSRKFFAQIVLPAVFVCIALVFSILVPPFKYPSLE
 LOPMYNEOYTFVNDAPEDTGTLELLNALTKDGGFTRCMENIPDTPCOAGREEW
 TTAPVPQTTMDLFONGNWTMNPSPACQSSDKIKKMLPVCPPGAGGLPPQPKONTA
 DILQDLTGNISDYLKVTYQIIAKSLNKNWNEFRYGGFSLGVSNTQALPPSOEVN
 DAIKMKHKLKAKSSADRFNLNSLGRMTQDLDRNNKVMWNNKGWHAISFLVNI
 NAILRANQKGENPSHYGTAFNHLNLTQOLSEVALMTTSVDLVSTICVIFAMSFV
 PASVFNLIQERVSRAKHQFISGVKPVYIWSNFVDMCNVVPATLVIIIFIGFOQ
 KSVVSTNPLVALLLLYGSITPLMPYSPAFVKIPSTAVVLTSVNLFTGINGSVA
 TVLELFTDNKNLNNILKVELIFPHCEGLRGLDMVKKQAMADALERGEFRFVS
 PLSWDLVGNLLEAVEGVVFLIIVLYQYFRIPRPVNAKLSPLNDEDEDVRRERQ
 RILDGGQNDLIEKLTIKYRRKRKPAVDRCVGPPECGFLGVNGAGKSSFTKM
 LTGDTVTGDAFLAKNSILNHEVHNMGYCQPODAITELLTGREHVFALLRGV
 PEKEVKGGEWAIRKGLVYKGYAGNYSGNKRLSTAMALIGGPPVFLDEPTTG
 MDPKARFLLNCALSVVKEGSRVLTSHSMECEALCTMAIMVNGRFRCLSGVQHLA
 NREGDGYTIVRIAGSNPDLKPVQDFGLAPGSLVEKHKRNLMQYQLPSSLSLARI
 FSLISQSKRLHIEDYSYSQTLQOVFNFAKQSDDDHLKDLSLHKNQTVVDVAULT
 SFLQEKVKESYV"
 72893..72986
 /gene="ABCA1"
 /number=3
 77520..77661
 /gene="ABCA1"
 /number=4
 78931..79049
 /gene="ABCA1"
 /number=5
 101048..101169
 /gene="ABCA1"
 /number=6
 104152..104328
 /gene="ABCA1"
 /number=7
 117287..117379
 /gene="ABCA1"
 /number=8
 123239..123569
 /gene="ABCA1"
 /number=9
 125285..125424
 /gene="ABCA1"
 /number=10
 125754..125870
 /gene="ABCA1"
 /number=11
 130079..130276
 /gene="ABCA1"
 /number=12
 131023..131228
 /gene="ABCA1"
 /number=13
 131749..131925
 /gene="ABCA1"
 /number=14
 133708..133930

```
/gene="ABCA1"
/number=15
135677..135898
/gene="ABCA1"
/number=16
136959..137163
/gene="ABCA1"
/number=17
138268..138381
/gene="ABCA1"
/number=18
140179..140350
/gene="ABCA1"
/number=19
141340..141471
/gene="ABCA1"
/number=20
142777..142919
/gene="ABCA1"
/number=21
143123..143260
/gene="ABCA1"
/number=22
143963..144183
/gene="ABCA1"
/number=23
145441..145513
/gene="ABCA1"
/number=24
146657..146859
/gene="ABCA1"
/number=25
148527..148575
/gene="ABCA1"
/number=26
148771..148884
/gene="ABCA1"
/number=27
150280..150428
/gene="ABCA1"
/number=28
152078..152202
/gene="ABCA1"
/number=29
153438..153536
/gene="ABCA1"
/number=30
156568..156757
/gene="ABCA1"
/number=31

Query Match 100.0%; Score 221; DB 9; Length 201144;
Best Local Similarity 100.0%; Pred. No. 5.5e-118; Mismatches 0; Indels 0; Gaps 0;
Matches 221; Conservative 0;

QY 1 gtaattgcgagcagagtagtgagtgaggccggaccgcagagccgagccgacccctctctc 60
Db 33931 GTAATTGCGAGCGAGAGTGAGTGGGCGCGGACCCGCGAGCGGACCGACCTTCTCTC 33990

QY 61 ccggggtcgcgagcagggcgaggctccgcgcacacacagagccgggtctcagggc 120
Db 33991 CCGGGTGGCGGACGAGCGGCGGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGCG 34050

QY 121 gctttgctctgtttttcccggttctgttttctcccttctccggaaggcttctcaa 180
Db 34051 GCITTTGCTCTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGGCTTGTCAA 34110

QY 181 ggggtagagaagagacgcaaacacaaagtggaaaacag 221
Db 34111 GGGGTAGAGAAAGAGACGCAACACAAAAGTGGAACACAG 34151

RESULT 11
```

```
AK022254
LOCUS Homo sapiens cDNA FLJ12192 fis, clone MAMMA1000851.
ACCESSION AK022254
VERSION AK022254.1 GI:10433612
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens Mammary gland cDNA to mRNA, clone_lib:MAMMAL
clone:MAMMA1000851.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Oshima,A.
NEDO human cDNA sequencing project
Unpublished (2000)
REFERENCE 2 (bases 1 to 1750)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
Kisarazu, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
Source 1..1750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1000851"
/clone_lib="MAMMAL"
/tissue_type="Mammary gland"
/note="Cloning vector: pME18SFL3"
BASE COUNT 291 a 489 c 586 g 384 t
ORIGIN

Query Match 99.1%; Score 219; DB 9; Length 1750;
Best Local Similarity 100.0%; Pred. No. 7.8e-117;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 aattgcgagcagagtagtgaggccggaccgcagagccgagccgacccctctctccc 62
Db 1 AATTGCGAGCGAGAGTGAGTGGGCGGACCCGCGAGAGCGCGACCCCTTCTCTCCC 60

QY 63 gggctcgcgagcagcagcggagagctccgcgcacacagagccggtcttcaggcgcc 122
Db 61 GGGCTCGCGAGCGGAGCGGCGGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGCGC 120

QY 123 ttgtctctgtttttcccggttctgttttctcccttctccggaaggcttgcagg 182
Db 121 TTGTCTCTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGGTTGTCAAG 180

QY 183 gtagagaaagacgacgcaaacacaaagtggaaaacag 221
Db 181 GGTAGAGAAAGAGACGCAACACAAAAGTGGAACACAG 219

RESULT 12
AK024328
LOCUS Homo sapiens cDNA FLJ14266 fis, clone PLACE1002437, highly similar
DEFINITION to ATP-BINDING CASSETTE TRANSPORTER 1.
ACCESSION AK024328
VERSION AK024328.1 GI:10436685
```


* sequencing reads tha

* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled.

* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 861 bp in length
* 55230 55329: gap of 100 bp
* 55330 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length

* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match 91.0%; Score 201; DB 2; Length 69570;
Best Local Similarity 100.0%; Pred. No. 2.8e-106;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgagcagagtgagtgaggccgggacccagagccgagccaccttctc 60
DB 41565 GTAATTGCCAGCGAGAGTGAGTGGGGCCGGACCCGACGAGCGCGACCTTCTCTC 41624
|||||
QY 61 ccgggctgcggcagcagcagcggcgagctccgcgcacacacagacccggttctcagggc 120
DB 41625 CCGGGCTGCGGACGAGCGGGCGGGAGCTCCGCGCACCAACAGAGCCGGTCTCAGGGC 41684
|||||
QY 121 gtttgcctctgtttttcccccgggtctgttttctccctctctccggaagcgttgtcaa 180
DB 41685 GCCTTGTCTCTGTTTTCCTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAA 41744
|||||
QY 181 ggggtaggagaaagacgca 201
DB 41745 GGGGTAGGAGAAAGAGACGCA 41765
|||||

RESULT 16
AX092589 AX092589 183999 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 1 from Patent WO0115676.
DEFINITION AX092589
ACCESSION AX092589
VERSION AX092589.1 GI:13444647
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 183999)
AUTHORS Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
TITLE Compositions and methods for modulating hdl cholesterol and triglyceride levels
JOURNAL Patent: WO 0115676-A 1 08-MAR-2001;
University of British Columbia (CA); Xenon Genetics Inc. (CA)
FEATURES
Source Location/Qualifiers
1..183999
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 49549 a 37944 c 41170 g 54950 t 386 others
ORIGIN

Query Match 91.0%; Score 201; DB 6; Length 183999;
Best Local Similarity 100.0%; Pred. No. 2.8e-106;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgagcagagtgagtgaggccgggacccagagccgagccaccttctc 60
DB 28790 GTAATTGCCAGCGAGAGTGAGTGGGGCCGGACCCGACGAGCGCGACCTTCTCTC 28849
|||||
QY 61 ccgggctgcggcagcagcagcggcgagctccgcgcacacacagacccggttctcagggc 120
DB 28850 CCGGGCTGCGGACGAGCGGGCGGGAGCTCCGCGCACCAACAGAGCCGGTCTCAGGGC 28909
|||||
QY 121 gtttgcctctgtttttcccccgggtctgttttctccctctctccggaagcgttgtcaa 180
DB 28910 GCTTGTCTCTGTTTTCCTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAA 28969
|||||
QY 181 ggggtaggagaaagacgca 201
DB 28970 GGGGTAGGAGAAAGAGACGCA 28990
|||||


```
RESULT 17
AX060713
LOCUS AX060713 10442 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 1 from Patent WO0078972.
ACCESSION AX060713
VERSION AX060713.1 GI:12406103
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn,R.M., Wade,D. and Garvin,M.
TITLE Regulation with binding cassette transporter protein abcl
JOURNAL Patent: WO 0078972-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
1..10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
ORIGIN
Query Match 89.1%; Score 197; DB 6; Length 10442;
Best Local Similarity 100.0%; Pred. No. 5.9e-104;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggacccgcagagccgagccgacctctctccgggctgcgagcgagcgagcg 84
|||||
Db 1 GCCCGGACCCGCGAGCCGAGCCGAGCCGACCCCTCTCTCCCGGCTCGGCGAGCGGCGG 60
|||||

QY 85 ggagctccgcgcacacacagagcggttctcagggcgttctgctctctctctctcc 144
|||||
Db 61 GGAGCTCCGGCGCACCAACAGAGCGGTTCTCAGGGGCTGTCTCAGGGGTAGGAGAAAGAGACGCAAC 120
|||||

QY 145 gttctgtttctccctctccggaagcgttctcaagggttaggagaaagacgcaaac 204
|||||
Db 121 GTTCTGTTTCTCCCTCTCTCCGGAAGCCTGTCTCAGGGGTAGGAGAAAGAGACGCAAC 180
|||||

QY 205 acaaaagtggaaacag 221
|||||
Db 181 ACAAAAGTGGAACAG 197
|||||

RESULT 18
AX060892
LOCUS AX060892 10442 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 1 from Patent WO0078971.
ACCESSION AX060892
VERSION AX060892.1 GI:12406270
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE Atp binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
1..10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
ORIGIN
Query Match 89.1%; Score 197; DB 6; Length 10442;
Best Local Similarity 100.0%; Pred. No. 5.9e-104;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggacccgcagagccgagccgacctctctccgggctgcgagcgagcgagcg 84
|||||
Db 1 GCCCGGACCCGCGAGCCGAGCCGAGCCGACCCCTCTCTCCCGGCTCGGCGAGCGGCGG 60
|||||

QY 85 ggagctccgcgcacacacagagcggttctcagggcgttctgctctctctctctcc 144
|||||
Db 61 GGAGCTCCGGCGCACCAACAGAGCGGTTCTCAGGGGCTGTCTCAGGGGTAGGAGAAAGAGACGCAAC 120
|||||

QY 145 gttctgtttctccctctccggaagcgttctcaagggttaggagaaagacgcaaac 204
|||||
Db 121 GTTCTGTTTCTCCCTCTCTCCGGAAGCCTGTCTCAGGGGTAGGAGAAAGAGACGCAAC 180
|||||

QY 205 acaaaagtggaaacag 221
|||||
Db 181 ACAAAAGTGGAACAG 197
|||||

RESULT 19
AX060713
LOCUS AF285167 10442 bp mRNA linear PRI 09-AUG-2000
DEFINITION Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA,
complete cds.
ACCESSION AF285167
VERSION AF285167.1 GI:9755158
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10442)
AUTHORS Schwartz,K., Lawn,R.M. and Wade,D.P.
TITLE ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
regulated by LXR
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 10442)
AUTHORS Lawn,R.M., Wade,D.P., Garvin,M.R., Wang,X., Schwartz,K.,
Porter,J.G., Seilhamer,J.J., Vaughan,A.M. and Oram,J.F.
TITLE Direct Submision
JOURNAL Submitted (06-JUL-2000) Discovery Research, CV Therapeutics Inc.,
3172 Porter Drive, Palo Alto, CA 94304, USA
FEATURES
Location/Qualifiers
1..10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/cell_type="fibroblast"
/tissue_type="skin"
1..10442
/gene="ABCA1"
291..7076
/gene="ABCA1"
/codon_start=1
/product="Atp-binding cassette transporter 1"
/protein_id="AAF98175.1"
/db_xref="GI:9755159"
/translation="MACWPQLRLLLKWLKLTERRQTCLLELVANPLFIFLILSVRL
SYPEYOEHECHPEKAMPAGTLPWQGLICNANPCFXYPTCPAGVGVGNFNKSLV
ARLFSARLLLYSQKTSMDMKVLRATLQIKSSNLKLODFLDVNETFSCFLVH
NLSLPKSTVDKMLRADYILHKVFLQGYQLHLTSLCNGSKSEEMQLQDQVESELCGLP
KEKLAARVLRNNMDILPLRLTINSTPFPKSLAEATKTLHLSTGLTAELEFSMR
SWDMROEVMFTNVNSSSTQIYQVSRVCGHPGEGGLKIKSLNMYEDNNYKALF
GNGTEDEATFYDNTPTPCNDMLNLESSPLRIIWKALKPLLVCKILVTPDTPAT
RQVAEYNKTEQLAVFHDLSGMHEELSPKIWTFNENSOENDLVRLMLSDNDHFWE
COLDLDTAQDIIAFLAKHEDKVOSSNGSVITWREAFNENQAIIRISRPMECNLN
KLEPATEWMLINKSMELLDERKFWAGIVFTGTPGSELPHHVKYKIRIMDINVERT
NKIKDGYWDPGRADPFEDMRYVMGFAYLQDVVEQAIIRVLTGRTKRTGYVMQMPY
PCYVDDIFLRVMSRSMPLMTLAWISVAVIKGIVYEKARLMTIRGLDMSILW
FSWPTSSILPLLVSAGLLVILKGNLLPYSDPSVVFVLSFAVAVTILQCLFLISTLF
SRANLAACAGGIIIFTLYLPLVLCVWQDYGFTLKIFASLLSPVAFGFGCEYFALPE
EQGIGVQNDNLFPESPEEDGNLTSTISMMLFDFTLYGVMTWYIEAVPPGYGIPRW
YFPCTKSYFGEESDEKSHPGSNQKRMSEICMEEPETHLKGIVSQNLVAVYRDMKV
AVDGLALNFYEGQITTSFLGHNGAGKTTTMSILTGLFPPTSGTAVILGDKIRSEMSSTR
```

QNLGVCPOHNVLFDMLTVEEHITWYARLKGSEKHVKAEMQMALDVLGSPSSKLSKT
 SOLSGMQRKLSALAFVGGSKVWILDEPTAGDVPYSRRGJWELLKYLKROGRTILST
 HMDADVLKIALISHGKCCVGSFLKNGLTGTYLTLVKKDVSSLSGCRNS
 STVSYLKEDSVSSQSDAGLSDHEDTLTIDVSAISNLRLKHXSEARLVEDIGHEL
 TVLYEAKGACAFVELBHEIDRSLDGLISYGSITTELEIFLLKVAESGVDVETS
 DDTLARRNRRAFQDKQCLRPFTEDDAADNDSDIDPESRETLKLSMGDGKSYQVK
 GMLTKQOQFVALLMKRLLTARRSRKFFAIVPAVFVIALVSLIIVPPFKYPSLE
 LQPMWNYEPTVMSNDAPEDTGTLELLNALTQDPGTRCMENP IPDTPCQAGEEW
 TTAPOFTIMDLFQNGWNTMNPSPACSSDKTKMLPVCPCGAGGLPPORONTA
 DILQDTGRNTSDYLKTYVOIIAISKLNKLTWNEFRYGGFSLGVSNTQALPPQEVN
 DAUKMKHKLAKDSSADRLNSUGREMTGLDTRNNVFNWFMHAISSFLIN
 NALFRANLQNGENSHYGTAFNPPLNTKQOLSEVALMTTSDVLSVICYIFAMSEV
 PASFVLTQIERSVYKLTGSGKPVLYISNFWMDMNVVPATVIVIFICFQ
 KSVSTVPLALLLLYGSITPLMPYPSAFVFKIPSTAVVLTSLNVLFGINGSVA
 TVLELFTDNKLNLDILKSVLFIHPHCLGRGLIDMKVKNQAMADALREGENRVS
 PLUSDVGRNIFAMAVEGVFLTLVLYOYRFFIPPPVNAKLSPLNDEDDVRRERQ
 RLDDGGNDLILEIKELKLYIRRRKKPAVDKICVGPCECFGLLVNGAGKSTFKM
 LTGDTVTRGDALFNKNSILNIEHVHNMGCPOFDATTELTGREGHEFFALLRG
 PKEVKYGEWNAIRKGLVYKGYAGYSGNKRKLSTAMALIGGPPVFLDPTTG
 MDPKARRELWNCALSVKGRSVLTSMSCEALCTRMAIMVNGFRCLGSLVOHLK
 NFRFGDYITVVRIGSNPDLPVODFGLAFPGSVLKEKHRLMLOYLPPSSLSIARI
 PSILSQRKRLHIEDYSYSQTTLDQVFNFAKQSDDDHKLDSLHKNQTVVDVAVLT
 SFLODEKVKESV"

BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
 ORIGIN

Query Match 89.1%; Score 197; DB 9; Length 10442;
 Best Local Similarity 100.0%; Pred. No. 5.9e-104;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggaccgagcagccgaccctctctccggcgctggcgagggcagggcg 84
 DB 1 GCGCGGAGCCGAGAGCGCGCGCCCTCTCTCCGGGCTGCGGAGGCGGCGG 60

QY 85 ggaqctccgcgcacacagccgctctcagggcgcttctctctctctctctcccg 144
 DB 61 GGAGCTCCCGCACACAGACCCGGTCTCAGGGCGCTTGTCTCTCTCTCTCTCTCCCG 120

QY 145 gttctgtttctccctctctccgagcgttctcaggggtaggagagagagacacac 204
 DB 121 GTTCTGTCTCTCTCTCTCTCCGAGGCTTGTCAAGGGTAGGAGAGAGACCAAC 180

QY 205 acaaaagtgaacacag 221
 DB 181 ACAAAGTGGAAACAG 197

RESULT 20
 AC021345/c
 LOCUS AC021345 90598 bp DNA linear HTG 13-JUL-2000
 DEFINITION Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC021345
 VERSION AC021345.2 GI:9130845
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 90598)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens, clone RP11-24J9
 Unpublished
 2 (bases 1 to 90598)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Becker, R., Beda, F.,
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
 Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
 Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenescor, J.,
 Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
 Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,

TITLE
 JOURNAL
 COMMENT

Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, F.M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severi, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6705761.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WfBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4483
 Center clone name: 24_J_9

* NOTE: This record contains 92 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 910: contig of 910 bp in length
 911 1010: gap of 100 bp
 1011 1873: contig of 863 bp in length
 1874 1973: gap of 100 bp
 1974 2824: contig of 851 bp in length
 2825 2924: gap of 100 bp
 2925 3802: contig of 878 bp in length
 3803 3902: gap of 100 bp
 3903 4816: contig of 914 bp in length
 4817 4916: gap of 100 bp
 4917 5759: contig of 843 bp in length
 5760 5859: gap of 100 bp
 5860 6784: contig of 905 bp in length
 6785 6864: gap of 100 bp
 6865 7747: contig of 883 bp in length
 7748 7847: gap of 100 bp
 7848 8755: contig of 908 bp in length
 8756 8855: gap of 100 bp
 8856 9733: contig of 898 bp in length
 9734 9833: gap of 100 bp
 9834 10757: contig of 904 bp in length
 10758 10857: gap of 100 bp
 10858 11732: contig of 875 bp in length
 11733 11832: gap of 100 bp
 11833 12739: contig of 907 bp in length
 12740 12839: gap of 100 bp
 12840 13710: contig of 871 bp in length
 13711 13810: gap of 100 bp
 13811 14684: contig of 874 bp in length
 14685 14784: gap of 100 bp
 14785 15662: contig of 878 bp in length
 15663 15762: gap of 100 bp
 15763 16677: contig of 915 bp in length
 16678 16777: gap of 100 bp
 16778 17678: contig of 901 bp in length
 17679 17778: gap of 100 bp
 17779 18679: contig of 901 bp in length
 18680 18779: gap of 100 bp


```
BASE COUNT      2906 a   2305 c   2416 g   2843 t   4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.le-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagcgagctccg 93
    |||||||
Db 42 CCGCAGAGCCGAGCCGACCTTCTCTCCGGGCTGCGGCGAGGCGAGGCGGGAGCTCCG 101
    |||||||

QY 94 cgcaccaacagagcggtctcagggcgcttgctcctgtttttcccggttctgttt 153
    |||||||
Db 102 CGCACCAACAGAGCGGTTCTCAGGGGCTTGCTCTCTGTTTTTCCCGGTTCTGTTT 161
    |||||||

QY 154 tctccctctccgaaagctgtcaagggtagagaaagagacgcaaacacaaaagt 213
    |||||||
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAAAGAGACGCAACACAAAAGTG 221
    |||||||

QY 214 gaaaacag 221
    |||||||
Db 222 GAAACAG 229

RESULT 22
AX060721
LOCUS      AX060721      10474 bp      DNA      linear      PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078972.
ACCESSION  AX060721
VERSION     AX060721.1 GI:12406109
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 10474)
JOURNAL    Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
CV         Regulation with binding cassette transporter protein abcl
PATENT: WO 0078972-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
SOURCE      1..10474
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      2907 a   2304 c   2415 g   2844 t   4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.le-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagcgagctccg 93
    |||||||
Db 42 CCGCAGAGCCGAGCCGACCTTCTCTCCGGGCTGCGGCGAGGCGAGGCGGGAGCTCCG 101
    |||||||

QY 94 cgcaccaacagagcggtctcagggcgcttgctcctgtttttcccggttctgttt 153
    |||||||
Db 102 CGCACCAACAGAGCGGTTCTCAGGGGCTTGCTCTCTGTTTTTCCCGGTTCTGTTT 161
    |||||||

QY 154 tctccctctccgaaagctgtcaagggtagagaaagagacgcaaacacaaaagt 213
    |||||||
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAAAGAGACGCAACACAAAAGTG 221
    |||||||

QY 214 gaaaacag 221
    |||||||
Db 222 GAAACAG 229

RESULT 23
AX060898
LOCUS      AX060898      10474 bp      DNA      linear      PAT 22-JAN-2001
DEFINITION Sequence 7 from Patent WO0078971.
ACCESSION  AX060898
VERSION     AX060898.1 GI:12406275
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 10474)
JOURNAL    Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
CV         Atp binding cassette transporter protein abcl polypeptides
PATENT: WO 0078971-A 7 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
SOURCE      1..10474
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT      2906 a   2305 c   2416 g   2843 t   4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.le-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagcgagctccg 93
    |||||||
Db 42 CCGCAGAGCCGAGCCGACCTTCTCTCCGGGCTGCGGCGAGGCGAGGCGGGAGCTCCG 101
    |||||||

QY 94 cgcaccaacagagcggtctcagggcgcttgctcctgtttttcccggttctgttt 153
    |||||||
Db 102 CGCACCAACAGAGCGGTTCTCAGGGGCTTGCTCTCTGTTTTTCCCGGTTCTGTTT 161
    |||||||

QY 154 tctccctctccgaaagctgtcaagggtagagaaagagacgcaaacacaaaagt 213
    |||||||
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAAAGAGACGCAACACAAAAGTG 221
    |||||||

QY 214 gaaaacag 221
    |||||||
Db 222 GAAACAG 229
```

```

QY 94 cgcacacacagcgcggttcctcagggccttctgtctcctgtttttcccccgttctgttt 153
Db 102 CGCACCAACAGAGCGGCTCTCAGGCGCTTGTGCTCTGTGTTTTTCCCGGTTCTGTGTT 161
QY 154 tctccctctcgcgaagcctgtcagggtagagaaagagacgcaaacacacaaagtq 213
Db 162 TCTCCCTCTCTCCGGAAGCTTGTCAAGGGTAGGAGAAAGAGACGCCAAACACAAAAGTG 221
QY 214 gaaaacag 221
Db 222 GAAACAG 229

RESULT 25
AX127764
LOCUS AX127764 446 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 3 from Patent WO0130848.
ACCESSION AX127764
VERSION AX127764.1 GI:14134411
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 446)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 3 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcctgtcgaagggtagga 189
Db 1 CTTGTTTTTCCCGGTTCTGTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60

QY 190 gaaagagacgcaaacacaaagtggaaaacag 221
Db 61 GAAAGAGACGCAAAACACAAAAGTGGAACAG 92

RESULT 26
AX139751
LOCUS AX139751 446 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1096012.
ACCESSION AX139751
VERSION AX139751.1 GI:14275333
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 446)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,I.I., G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcctgtcgaagggtagga 189
Db 1 CTTGTTTTTCCCGGTTCTGTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60

QY 190 gaaagagacgcaaacacaaagtggaaaacag 221
Db 61 GAAAGAGACGCAAAACACAAAAGTGGAACAG 92

```

```

source 1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcctgtcgaagggtagga 189
Db 1 CTTGTTTTTCCCGGTTCTGTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60

QY 190 gaaagagacgcaaacacaaagtggaaaacag 221
Db 61 GAAAGAGACGCAAAACACAAAAGTGGAACAG 92

RESULT 27
AX127830
LOCUS AX127830 9741 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 69 from Patent WO0130848.
ACCESSION AX127830
VERSION AX127830.1 GI:14134477
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9741)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 69 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcctgtcgaagggtagga 189
Db 1 CTTGTTTTTCCCGGTTCTGTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60

QY 190 gaaagagacgcaaacacaaagtggaaaacag 221
Db 61 GAAAGAGACGCAAAACACAAAAGTGGAACAG 92

RESULT 28
AX139817
LOCUS AX139817 9741 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 69 from Patent EP1096012.
ACCESSION AX139817
VERSION AX139817.1 GI:14275399
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9741)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcctgtcgaagggtagga 189
Db 1 CTTGTTTTTCCCGGTTCTGTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60

QY 190 gaaagagacgcaaacacaaagtggaaaacag 221
Db 61 GAAAGAGACGCAAAACACAAAAGTGGAACAG 92

```

AUTHORS Denefle, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C., Naudin, L., Lemoine, C., Duverger, N., Jaye, M., searfoos Iii, G.H., Remaley, A., Brewer, H.B. and Dean, M.

TITLE Nucleic acids of the human abcl gene and their therapeutic and diagnostic application

JOURNAL Patent: EP 1056012-A 69 02-MAY-2001; Aventis Pharma S.A. (FR)

FEATURES
source Location/Qualifiers
1. .9741
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others

ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccggttctgttttcccttcccttccggaagctgttgaagggttagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGAGAGCTTGTCAAGGGTAGGA 60
|||||

QY 190 gaaagagcgcacacacaaaagtgaacaacag 221
|||||

Db 61 GAAAGAGCGCAACACAAAAGTGGAACACAG 92

RESULT 29
AX351038 9741 bp DNA linear PAT 06-FEB-2002

LOCUS AX351038

DEFINITION Sequence 10 from Patent WO0183746.

ACCESSION AX351038

VERSION AX351038.1 GI:18616393

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Pojo, S.

TITLE Regulatory nucleic acid sequences of the abcl gene

JOURNAL Patent: WO 0183746-A 10 08-NOV-2001; Aventis Pharma S.A. (FR)

FEATURES
source Location/Qualifiers
1. .9741
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others

ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccggttctgttttcccttcccttccggaagctgttgaagggttagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGAGAGCTTGTCAAGGGTAGGA 60
|||||

QY 190 gaaagagcgcacacacaaaagtgaacaacag 221
|||||

Db 61 GAAAGAGCGCAACACAAAAGTGGAACACAG 92

RESULT 30
AX060715 1643 bp DNA linear PAT 22-JAN-2001

LOCUS AX060715

DEFINITION Sequence 3 from Patent WO0078972.

ACCESSION AX060715

VERSION AX060715.1 GI:12406104

KEYWORDS human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Pojo, S.

TITLE Regulatory nucleic acid sequences of the abcl gene

JOURNAL Patent: WO 0183746-A 10 08-NOV-2001; Aventis Pharma S.A. (FR)

FEATURES
source Location/Qualifiers
1. .9741
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others

ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccggttctgttttcccttcccttccggaagctgttgaagggttagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGAGAGCTTGTCAAGGGTAGGA 60
|||||

QY 190 gaaagagcgcacacacaaaagtgaacaacag 221
|||||

Db 61 GAAAGAGCGCAACACAAAAGTGGAACACAG 92

RESULT 30
AX060715 1643 bp DNA linear PAT 22-JAN-2001

LOCUS AX060715

DEFINITION Sequence 3 from Patent WO0078972.

ACCESSION AX060715

VERSION AX060715.1 GI:12406104

KEYWORDS human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Pojo, S.

TITLE Regulatory nucleic acid sequences of the abcl gene

JOURNAL Patent: WO 0183746-A 10 08-NOV-2001; Aventis Pharma S.A. (FR)

FEATURES
source Location/Qualifiers
1. .9741
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others

ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccggttctgttttcccttcccttccggaagctgttgaagggttagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGAGAGCTTGTCAAGGGTAGGA 60
|||||

QY 190 gaaagagcgcacacacaaaagtgaacaacag 221
|||||

Db 61 GAAAGAGCGCAACACAAAAGTGGAACACAG 92

SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS 1 (bases 1 to 1643)
Lawn, R.M., Wade, D. and Garvin, M.

TITLE Regulation with binding cassette transporter protein abcl

JOURNAL Patent: WO 0078972-A 3 28-DEC-2000; CV THERAPEUTICS, INC. (US)

FEATURES
source Location/Qualifiers
1. .1643
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 370 a 413 c 457 g 403 t

ORIGIN

Query Match 41.2%; Score 91; DB 6; Length 1643;
Best Local Similarity 100.0%; Pred. No. 6.4e-42;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgacgagagtgagtgccgggacccgcagagccgagccaccccttcttc 60
|||||

Db 1553 GTAATTGCGAGCAGAGTGAGTGGCGCGGACCCGAGAGCCGAGCCGACCTTCTCTC 1612
|||||

QY 61 ccgggctgcgcagggcagggcagggcgagctc 91
|||||

Db 1613 CCGGGCTGCGGCGAGGCGAGGCGGGGAGCTC 1643

RESULT 31
AX060894 1643 bp DNA linear PAT 22-JAN-2001

LOCUS AX060894

DEFINITION Sequence 3 from Patent WO0078971.

ACCESSION AX060894

VERSION AX060894.1 GI:12406271

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS 1 (bases 1 to 1643)
Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.

TITLE Atp binding cassette transporter protein abcl polypeptides

JOURNAL Patent: WO 0078971-A 3 28-DEC-2000; CV THERAPEUTICS, INC. (US)

FEATURES
source Location/Qualifiers
1. .1643
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 370 a 413 c 457 g 403 t

ORIGIN

Query Match 41.2%; Score 91; DB 6; Length 1643;
Best Local Similarity 100.0%; Pred. No. 6.4e-42;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgacgagagtgagtgccgggacccgcagagccgagccaccccttcttc 60
|||||

Db 1553 GTAATTGCGAGCAGAGTGAGTGGCGCGGACCCGAGAGCCGAGCCGACCTTCTCTC 1612
|||||

QY 61 ccgggctgcgcagggcagggcagggcgagctc 91
|||||

Db 1613 CCGGGCTGCGGCGAGGCGAGGCGGGGAGCTC 1643

RESULT 32
AC021246/c 69570 bp DNA linear HTG 13-JUL-2000

LOCUS AC021246

DEFINITION Homo sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.

ACCESSION AC021246

VERSION AC021246.2 GI:9119882

KEYWORDS HTG: HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome, clone RP11-1N10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 69570)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeRrellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, W.

TITLE Direct Submission
JOURNAL Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 13, 2000 this sequence version replaced gl:6705871.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997).
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: L_N_10

* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 871: contig of 871 bp in length
* 872 971: gap of 100 bp
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length
* 3746 3845: gap of 100 bp
* 3846 4696: contig of 851 bp in length
* 4697 4796: gap of 100 bp
* 4797 5640: contig of 844 bp in length
* 5641 5740: gap of 100 bp
* 5741 6540: contig of 800 bp in length
* 6541 6640: gap of 100 bp
* 6641 7509: contig of 869 bp in length
* 7510 7609: gap of 100 bp
* 7610 8479: contig of 870 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 880 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length
* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length
* 20922 21021: gap of 100 bp
* 21022 21865: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 26621: contig of 851 bp in length
* 26622 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 27676: gap of 100 bp
* 27677 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp

```

* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 861 bp in length
* 55230 55329: gap of 100 bp
* 55330 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match      21.3%; Score 47; DB 2; Length 69570;
Best Local Similarity 100.0%; Pred. No. 3.7e-16;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 caggcgaggagctccgcgcaccacagagccggtctcaggcgct 123
|||||
Db 67335 CAGGCGGGAGGCTCCCGCCACCAAGAGCCGGTTCTCAGGCGGCT 67289

RESULT 33
AC097517/c
LOCUS      AC097517      92227 bp      DNA      linear      PRI 03-JAN-2002
DEFINITION Homo sapiens chromosome 2 clone RP11-485017, complete sequence.
ACCESSION  AC097517
VERSION     AC097517.3  GI:117998642
KEYWORDS   HTG.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 92227)
AUTHORS   Waterston, R.H.

The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 92227)
Waterston, R.H.
Direct Submission
Submitted (18-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
3 (bases 1 to 92227)
Waterston, R.H.
Direct Submission
Submitted (29-DEC-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 92227)
Waterston, R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Dec 29, 2001 this sequence version replaced gi:16902004.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H_NH0485017
Drafting center: WIBR
-----
Location/Qualifiers
1..92227
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-485017"
BASE COUNT  27775 a 18464 c 18146 g 27842 t
ORIGIN

Query Match      9.0%; Score 20; DB 9; Length 92227;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 tgttttccctctctccgg 168
|||||
Db 73615 TGTGTTCTCCCTCTCTCCGG 73596

RESULT 34
AC022563
LOCUS      AC022563      149710 bp      DNA      linear      HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-3L23, LOW-PASS SEQUENCE SAMPLING.
ACCESSION  AC022563
VERSION     AC022563.2  GI:9121084
KEYWORDS   HTG; HTGS_PHASE0.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 149710)
AUTHORS   Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE     Homo sapiens chromosome, clone RP11-3L23
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 149710)
AUTHORS   Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
            Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
            Boguslavsky, I., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
            Choepel, I., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
            Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
            Ferrelira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
            Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,

```


Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meidrim, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., Olivat, T.M., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
Zimmer, A., and Zody, N.

TITLE JOURNAL

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6910806.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2771

Center clone name: 3_L_23

* NOTE: This record contains 171 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 682: contig of 682 bp in length
683 782: gap of 100 bp
783 1492: contig of 710 bp in length
1493 1592: gap of 100 bp
1593 2305: contig of 713 bp in length
2306 2405: gap of 100 bp
2406 3113: contig of 708 bp in length
3114 3213: gap of 100 bp
3214 3867: contig of 654 bp in length
3868 3967: gap of 100 bp
3968 4646: contig of 679 bp in length
4647 4746: gap of 100 bp
4747 5416: contig of 670 bp in length
5417 5516: gap of 100 bp
5517 6193: contig of 677 bp in length
6194 6293: gap of 100 bp
6294 6968: contig of 675 bp in length
6969 7068: gap of 100 bp
7069 7757: contig of 689 bp in length
7758 7857: gap of 100 bp
7858 8557: contig of 700 bp in length
8558 8657: gap of 100 bp
8658 9336: contig of 679 bp in length
9337 9436: gap of 100 bp
9437 10124: contig of 688 bp in length
10125 10224: gap of 100 bp
10225 10879: contig of 655 bp in length
10880 10979: gap of 100 bp
10980 11663: contig of 684 bp in length
11664 11763: gap of 100 bp
11764 12433: contig of 670 bp in length
12434 12533: gap of 100 bp
12534 13209: contig of 676 bp in length
13210 13309: gap of 100 bp
13310 13989: contig of 680 bp in length
13990 14089: gap of 100 bp

14090 14798: contig of 709 bp in length
14799 14898: gap of 100 bp
14899 15571: contig of 673 bp in length
15572 15671: gap of 100 bp
15672 16387: contig of 716 bp in length
16388 16487: gap of 100 bp
16488 17198: contig of 711 bp in length
17199 17298: gap of 100 bp
17299 18008: contig of 710 bp in length
18009 18108: gap of 100 bp
18109 18772: contig of 664 bp in length
18773 18872: gap of 100 bp
18873 19555: contig of 683 bp in length
19556 19655: gap of 100 bp
19656 20320: contig of 665 bp in length
20321 20420: gap of 100 bp
20421 21112: contig of 692 bp in length
21113 21212: gap of 100 bp
21213 21885: contig of 673 bp in length
21886 21985: gap of 100 bp
21986 22671: contig of 686 bp in length
22672 22771: gap of 100 bp
22772 23451: contig of 680 bp in length
23452 23551: gap of 100 bp
23552 24267: contig of 716 bp in length
24268 24367: gap of 100 bp
24368 25073: contig of 706 bp in length
25074 25173: gap of 100 bp
25174 25867: contig of 694 bp in length
25868 25967: gap of 100 bp
25968 26669: contig of 702 bp in length
26670 26769: gap of 100 bp
26770 27445: contig of 676 bp in length
27446 27545: gap of 100 bp
27546 28224: contig of 679 bp in length
28225 28324: gap of 100 bp
28325 29002: contig of 678 bp in length
29003 29102: gap of 100 bp
29103 29785: contig of 683 bp in length
29786 29885: gap of 100 bp
29886 30571: contig of 686 bp in length
30572 30671: gap of 100 bp
30672 31357: contig of 686 bp in length
31358 31457: gap of 100 bp
31458 32168: contig of 711 bp in length
32169 32268: gap of 100 bp
32269 32951: contig of 683 bp in length
32952 33051: gap of 100 bp
33052 33756: contig of 705 bp in length
33757 33856: gap of 100 bp
33857 34566: contig of 710 bp in length
34567 34666: gap of 100 bp
34667 35376: contig of 710 bp in length
35377 35476: gap of 100 bp
35477 36143: contig of 667 bp in length
36144 36243: gap of 100 bp
36244 36924: contig of 681 bp in length
36925 37024: gap of 100 bp
37025 37701: contig of 677 bp in length
37702 37801: gap of 100 bp
37802 38493: contig of 692 bp in length
38494 38593: gap of 100 bp
38594 39274: contig of 681 bp in length
39275 39374: gap of 100 bp
39375 40071: contig of 697 bp in length
40072 40171: gap of 100 bp
40172 40870: contig of 699 bp in length
40871 40970: gap of 100 bp
40971 41659: contig of 689 bp in length
41660 41759: gap of 100 bp
41760 42434: contig of 675 bp in length
42435 42534: gap of 100 bp
42535 43207: contig of 673 bp in length

* 43208 43307: gap of 100 bp
 * 43308 43980: contig of 673 bp in length
 * 43981 44080: gap of 100 bp
 * 44081 44768: contig of 688 bp in length
 * 44769 44868: gap of 100 bp
 * 44869 45556: contig of 688 bp in length
 * 45557 45656: gap of 100 bp
 * 45657 46348: contig of 692 bp in length
 * 46349 46448: gap of 100 bp
 * 46449 47143: contig of 695 bp in length
 * 47144 47243: gap of 100 bp
 * 47244 47943: contig of 700 bp in length
 * 47944 48043: gap of 100 bp
 * 48044 48727: contig of 684 bp in length
 * 48728 48827: gap of 100 bp
 * 48828 49540: contig of 713 bp in length
 * 49541 49640: gap of 100 bp
 * 49641 50332: contig of 692 bp in length
 * 50333 50432: gap of 100 bp
 * 50433 51127: contig of 695 bp in length
 * 51128 51227: gap of 100 bp
 * 51228 51915: contig of 688 bp in length
 * 51916 52015: gap of 100 bp
 * 52016 52678: contig of 663 bp in length
 * 52679 52778: gap of 100 bp
 * 52779 53460: contig of 682 bp in length
 * 53461 53560: gap of 100 bp
 * 53561 54251: contig of 691 bp in length
 * 54252 54351: gap of 100 bp
 * 54352 54921: contig of 570 bp in length
 * 54922 55021: gap of 100 bp
 * 55022 55702: contig of 681 bp in length
 * 55703 55802: gap of 100 bp

Query Match 9.0%; Score 20; DB 2; Length 149710;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 agcgtgtcgaagggttagga 189
 |||||
 Db 55886 AGGCTTCAAGGGTAGGA 55905

RESULT 35
 AC027453
 LOCUS Homo sapiens chromosome 10 clone RP11-524H12 map 10, WORKING DRAFT
 DEFINITION AC027453 162496 bp DNA linear HTG 26-MAY-2000
 ACCESSION AC027453
 VERSION AC027453.3 GI:8077127
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 162496)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 10, clone RP11-524H12
 Unpublished
 2 (bases 1 to 162496)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F.,
 Boguski,M.S., Bouckgeer,B., Brown,A., Burkett,G.,
 Campopiano,A., Cooke,A., Choehel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., DeAfrillano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Kiehl,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 25, 2000 this sequence version replaced gi:7637299.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L8240
 Center clone name: 524_H12

----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731

Consensus quality: 154128 bases at least Q40
 Consensus quality: 158489 bases at least Q30

Consensus quality: 160079 bases at least Q20
 Insert size: 163000; agarose-fp

Insert size: 160996; sum-of-ctgifs
 Quality coverage: 4.7 in Q20 bases; agarose-fp

Quality coverage: 4.8 in Q20 bases; sum-of-ctgifs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1
 * 882 981: gap of 100 bp
 * 982 4366: contig of 3385 bp in length
 * 4367 4466: gap of 100 bp
 * 4467 7420: contig of 2954 bp in length
 * 7421 7520: gap of 100 bp
 * 7521 11683: contig of 4163 bp in length
 * 11684 11783: gap of 100 bp
 * 11784 15264: contig of 3481 bp in length
 * 15265 15364: gap of 100 bp
 * 15365 20920: contig of 5556 bp in length
 * 20921 21020: gap of 100 bp
 * 21021 27969: contig of 6949 bp in length
 * 27970 28069: gap of 100 bp
 * 28070 36585: contig of 8516 bp in length
 * 36586 36685: gap of 100 bp
 * 36686 47838: contig of 11153 bp in length
 * 47839 47938: gap of 100 bp
 * 47939 57920: contig of 9982 bp in length
 * 57921 58020: gap of 100 bp
 * 58021 70481: contig of 12461 bp in length
 * 70482 70581: gap of 100 bp
 * 70582 80819: contig of 10238 bp in length
 * 80820 80919: gap of 100 bp
 * 80920 91740: contig of 10821 bp in length
 * 91741 91840: gap of 100 bp
 * 91841 113943: contig of 22103 bp in length
 * 113944 114043: gap of 100 bp
 * 114044 137119: contig of 23076 bp in length
 * 137120 137219: gap of 100 bp

TITLE JOURNAL

COMMENT

* 137220 162496: contig of 25277 bp in length.

FEATURES

source
1. .162496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10"
/clone="RP11-524H12"
/clone_lib="RPC1-11 Human Male BAC"
1. .881
/note="assembly_fragment
vector_end:T7
vector_side:left"
misc_feature 982. .4366
/note="assembly_fragment"
misc_feature 4467. .7420
/note="assembly_fragment"
misc_feature 7521. .11683
/note="assembly_fragment"
misc_feature 11784. .15264
/note="assembly_fragment
clone_end:SP6
vector_side:right"
misc_feature 15365. .20920
/note="assembly_fragment"
misc_feature 21021. .27969
/note="assembly_fragment"
misc_feature 28070. .36585
/note="assembly_fragment"
misc_feature 36686. .47838
/note="assembly_fragment"
misc_feature 47939. .57920
/note="assembly_fragment"
misc_feature 58021. .70481
/note="assembly_fragment"
misc_feature 70582. .80819
/note="assembly_fragment"
misc_feature 80920. .91740
/note="assembly_fragment"
misc_feature 91841. .113943
/note="assembly_fragment"
misc_feature 114044. .137119
/note="assembly_fragment"
misc_feature 137220. .162496
/note="assembly_fragment"
BASE COUNT 47419 a 33487 c 33493 g 46593 t 1504 others
ORIGIN

Query Match 9.0%; Score 20; DB 2; Length 162496;

Best Local Similarity 100.0%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;

Qy 149 tgttttcccttctccg 168

Db 113242 TGTTCCTCCCTCTCCG 113261

RESULT 36

AC012325 201938 bp DNA linear HTG 25-APR-2001
LOCUS Homo sapiens chromosome 16 clone RP11-93H5, WORKING DRAFT SEQUENCE,
DEFINITION 7 ordered pieces.

ACCESSION AC012325

VERSION AC012325.6 GI:9954638

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 201938)

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 16

JOURNAL

REFERENCE 2 (bases 1 to 201938)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL

Submitted (23-Oct-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Aug 31, 2000 this sequence version replaced gi:7690207.

COMMENT

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced by the finished sequence as soon as it is available and * the accession number will be preserved.

* 1 33343: contig of 33343 bp in length
* 33344 33443: gap of unknown length
* 33444 40628: contig of 7185 bp in length
* 40629 40728: gap of unknown length
* 40729 143349: contig of 102621 bp in length
* 143350 143449: gap of unknown length
* 143450 144499: contig of 1050 bp in length
* 144500 144599: gap of unknown length
* 144600 162264: contig of 17665 bp in length
* 162265 162364: gap of unknown length
* 162365 187856: contig of 25492 bp in length
* 187857 187956: gap of unknown length
* 187957 201938: contig of 13982 bp in length.

FEATURES

source

1..201938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-93H5"

BASE COUNT 54511 a 50297 c 47816 g 48714 t 600 others
ORIGIN

Query Match

Best Local Similarity 9.0%; Score 20; DB 2; Length 201938;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 gggctgcggcaggcaggc 82

Db 186942 GGGCTGCGCAGGCAGGC 186961

RESULT 37

AC097253/c 228703 bp DNA linear HTG 20-DEC-2001
LOCUS Rattus norvegicus chromosome SA clone CH230-152G15, WORKING DRAFT
DEFINITION SEQUENCE, 12 unordered pieces.

ACCESSION AC097253 GI:17973859
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 228703)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buha,Y., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Channon,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dunn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holliday,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lousleged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenwo,S., Oghu,M., Okwono,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,J., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

Direct Submission
 Unpublished
 2 (bases 1 to 228703)
 Worley,K.C.
 Direct Submission
 Submitted (13-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:17062530.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GIOR
 Center clone name: CH230-152G15
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329First call to findPhraplist
 Consensus quality: 221817 bases at least Q40
 Consensus quality: 223045 bases at least Q30
 Consensus quality: 223940 bases at least Q20
 Estimated insert size: 224711; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-ip estimation

 Quality coverage: 6.4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 40920: contig of 40920 bp in length
 * 40921 41020: gap of unknown length
 * 41021 73829: contig of 32809 bp in length
 * 73830 73929: gap of unknown length
 * 73930 100142: contig of 26213 bp in length
 * 100143 100242: gap of unknown length
 * 100243 126981: contig of 26739 bp in length
 * 126982 127081: gap of unknown length
 * 127082 150119: contig of 23038 bp in length
 * 150120 150219: gap of unknown length
 * 150220 172428: contig of 22209 bp in length
 * 172429 172528: gap of unknown length
 * 172529 191512: contig of 18983 bp in length
 * 191513 209925: contig of 18314 bp in length
 * 209926 210025: gap of unknown length
 * 210026 219870: contig of 9845 bp in length
 * 219871 225401: contig of 5431 bp in length
 * 225402 225501: gap of unknown length
 * 225502 227014: contig of 1513 bp in length
 * 227015 227114: gap of unknown length
 * 227115 228703: contig of 1589 bp in length.

FEATURES
 Location/Qualifiers
 1..228703
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /chromosome="SA"
 /clone="CH230-152G15"
 BASE COUNT 65603 a 46098 c 46478 g 69392 t 1132 others
 ORIGIN

Query Match 9.0%; Score 20; DB 2; Length 228703;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ttctgtttttctccctctc 165
 ||||||||||||||||
 DB 82733 TTCTGTTTTCCTCCCTCTC 82714

RESULT 38
 A39868/c
 LOCUS A39868
 DEFINITION Sequence 2 from Patent EP0614981.
 ACCESSION A39868
 VERSION A39868.1 GI:2296098
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 735)
 AUTHORS Dembic,Z., Garotta,G. and Gentz,R.H.
 TITLE Chimeric human interferon-gamma-receptor/immunoglobulin polypeptides
 JOURNAL Hoffmann LA ROCHE (CH)
 COMMENT Patent: EP 0614981-A 2 14-SEP-1994;
 Other publication JP 6319552 941122
 Other publication NZ 250997 951026
 Other publication CA 2114168 940906

Other publication CN 1094092 941026
Other publication AU 5647894 940908
Other publication ZA 9401333 940906.

FEATURES

Location/Qualifiers
1..735

/organism="unidentified"

/db_xref="taxon:32644"

223 a 135 c 170 g 207 t

BASE COUNT
ORIGIN

Query Match 8.6%; Score 19; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caagggttaggagaagag 196

|||||

Db 28 CAAGGGGTAGGAGAAAG 10

RESULT 39

AF056979/c

LOCUS

DEFINITION Homo sapiens clone YAN1 interferon-gamma receptor mRNA, complete cds.

ACCESSION AF056979

VERSION AF056979.1 GI:13562048

KEYWORDS

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Allende, L.M., Lopez-Goyanes, A., Paz-Artal, E., Correll, A.,

Garcia-Perez, M.A., Varela, P., Scarpellini, A., Negreira, S.,

Palenque, E. and Arnaiz-Villena, A.

A point mutation in a domain of gamma interferon receptor 1

provokes severe immunodeficiency

Clin. Diagn. Lab. Immunol. 8 (1), 133-137 (2001)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submited (02-APR-1998) Immunology, Hospital 12 de Octubre, Crta.

Andalucia Km. 5.4, Madrid 28041, Spain

Location/Qualifiers

1..1641

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="YAN1"

/sex="female"

/cell_line="HVS transformed"

/cell_type="lymphocyte"

/tissue_type="blood"

22..1491

/codon_start=1

/product="interferon-gamma receptor"

/protein_id="AAK30623.1"

/db_xref="GI:13562049"

/translation="MALLFLPLVMQGVSRAMGTADLGPSSVPTPTNTVTIESNMNP

IRVYQYQMPQVPVFTVEGKNGYKNSEWIDACINISHHCNIDSHVGDPSNLSWVR

KAVGQKESAPKFEFACRGDKIGPKPLDIRKEEKOIMIDIFHPSVFVNGDEQVD

YDPETTCIRVNYVYRMNGSEIQYKILTKQEDCDELOCOLAIPVSSLNSQYCVSAE

GLVHWGVTLSKREVCITIFNSSKSGSLWIPVVAALLFLVLSLFCIFKIKINPL

REKSIILPKSLSVYSATLETKPKSKYSLITSYQPSLEKVEVCEPLSPATVPM

HTEDNPGVHEELSSITEVTTENIPDPVPGSHLPIERSSPLSSNQSPFSI

ALNSVHSRNCSDHSRNGFDTDSCLSLSLSDSEPPNNKGEIKTEGQELITVIK

APTSPGYDKPHVLDLLVDDSGKESLIGRYTDSKEFS"

508 a 320 c 352 g 461 t

BASE COUNT

ORIGIN

Query Match 8.6%; Score 19; DB 9; Length 1641;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caagggttaggagaagag 196

|||||

Db 49 CAAGGGGTAGGAGAAAGAG 31

RESULT 40

SIVAGMGAGD

LOCUS

DEFINITION

Simian immunodeficiency virus (clone agm49gag) gag gene, complete

exon.

ACCESSION

LI9253

VERSION

LI9253.1 GI:310428

KEYWORDS

gag gene.

SOURCE

Simian immunodeficiency virus DNA.

ORGANISM

Viruses; Retroviridae; Retroviridae; Lentivirus; Primate

lentivirus group.

REFERENCE

AUTHORS

Hirsch, V.M., McGann, C., Dapolito, G., Goldstein, S., Ogen-Odoi, A.,

Biryawaho, B., Lakwo, T. and Johnson, P.R.

Identification of a new subgroup of SIVagm in tantalus monkeys

Virolgy 197, 426-430 (1993)

JOURNAL

MEDLINE

94025594

FEATURES

Location/Qualifiers

1..1773

/organism="Simian immunodeficiency virus"

/proviral

/db_xref="taxon:11723"

/clone="agm49gag"

/haplotype="na"

/cell_line="CEMss"

195..1773

/gene="gag"

195..1773

/gene="gag"

594 a 375 c 480 g 324 t

BASE COUNT

ORIGIN

Query Match 8.6%; Score 19; DB 14; Length 1773;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 gcttgtaggggttaggag 190

|||||

Db 1265 GCTTGTCAGGGGTAGGAG 1283

RESULT 41

HUMIFNRG/c

LOCUS

DEFINITION

Human interferon-gamma receptor mRNA, complete cds.

ACCESSION

J031143

VERSION

J031143.1 GI:184650

KEYWORDS

interferon receptor.

SOURCE

Human lymphoid tissue

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Aguet, M., Dembic, Z. and Merlin, G.

Molecular cloning and expression of the human interferon-gamma

receptor

Cell 55 (2), 273-280 (1988)

JOURNAL

MEDLINE

89003065

COMMENT

Draft entry and computer-readable sequence for [1] kindly provided

by M.Aguet, 08-SEP-1988.

Location/Qualifiers

1..2064

FEATURES

source

passed the following selection criteria: matched mRNA gi: 5419821.

FEATURES
source
Location/Qualifiers
1. .2127
/organism="Homo sapiens"
/db_xref="LocusID:3459"
/db_xref="taxon:9606"
/clone="MGC:12420 IMAGE:3950528"
/tissue_type="prostate"
/clone_lib="NIH_MGC_83"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
83. .1552
/codon_start=1
/product="interferon gamma receptor 1"
/protein_id="AAH05333.1"
/db_xref="GI:13529119"
/translation="MALLFLPLVMQGVSRAMGTADLGPSVPTTNTVIESYNNMP
IVWEYQIMQVPVFTVEVKNYGVKNSWDACINISHHYCNISDRHVGDPSPNSLWVRV
KARVGQESAYAKSEFAVCRDGIKPPKLDIRKEKQIMIDIFHPSVFNQDEQVD
YDPETTCYIRVYVYRMNGSEIQKILTKQEDDCDEIQCOLAIPVSSNSQYCVSAE
GVLHWGVTTERKSEVCITIFNSIKGSLWIPVVAALLFLVLSLFCFYIKKINPL
KEKSIILPKSLISVRSATLETPEKSYSLITSYQFSLKEVCEELSPATVPGM
HTEDNPKVTEELSSITEVTEENIPDVPGSHLTPIERESSPLSSNQSEPGSI
ALNSYHSRNCSESDSRNGFDTSCLSHSLSDSEFPNPKGEIKTEQEELITVVK
APTSFGYDKPHVLVDLLVDDSGKESLIGYRPTEDSKEFS"
BASE COUNT 671 a 391 c 445 g 620 t
ORIGIN

CDS

Query Match 8.6%; Score 19; DB 9; Length 2127;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 caaggggtaggagaaag 196
|||||
Db 110 CAAGGGTAGGAGAAAG 92

RESULT 43
LOCUS A30438 2184 bp DNA linear PAT 06-AUG-1996
DEFINITION H.sapiens pBABE Sact/Asp181 fragment.
ACCESSION A30438
VERSION A30438.1 GI:1567031
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2184)
AUTHORS Fountoulakis M., Garotta G. and Stueber D.
TITLE Soluble interferon-gamma receptors and methods for their production
JOURNAL Patent: EP 0393502-A 1 24-OCT-1990;
F. HOFFMANN-LA ROCHE AG

FEATURES
source
Location/Qualifiers
1. .2184
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 688 a 413 c 452 g 631 t
ORIGIN

Query Match 8.6%; Score 19; DB 6; Length 2184;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 caaggggtaggagaaag 196
|||||
Db 112 CAAGGGTAGGAGAAAG 94

RESULT 44
STH022

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="6q23-q24"
<1. .2064
/gene="IFNGR1"
/product="IFNR-gamma mRNA"
1. .2064
/gene="IFNGR1"
49. .1518
/gene="IFNGR1"
/note="interferon-gamma receptor"
/codon_start=1
/protein_id="AA52731.1"
/db_xref="GI:306915"
/db_xref="GDB:G00-120-688"
/translation="MALLFLPLVMQGVSRAMGTADLGPSVPTTNTVIESYNNMP
IVWEYQIMQVPVFTVEVKNYGVKNSWDACINISHHYCNISDRHVGDPSPNSLWVRV
KARVGQESAYAKSEFAVCRDGIKPPKLDIRKEKQIMIDIFHPSVFNQDEQVD
YDPETTCYIRVYVYRMNGSEIQKILTKQEDDCDEIQCOLAIPVSSNSQYCVSAE
GVLHWGVTTERKSEVCITIFNSIKGSLWIPVVAALLFLVLSLFCFYIKKINPL
KEKSIILPKSLISVRSATLETPEKSYSLITSYQFSLKEVCEELSPATVPGM
HTEDNPKVTEELSSITEVTEENIPDVPGSHLTPIERESSPLSSNQSEPGSI
ALNSYHSRNCSESDSRNGFDTSCLSHSLSDSEFPNPKGEIKTEQEELITVVK
APTSFGYDKPHVLVDLLVDDSGKESLIGYRPTEDSKEFS"
BASE COUNT 639 a 383 c 426 g 616 t
ORIGIN 1 bp upstream of EcoRI site; chromosome 6q15-q21.

Query Match 8.6%; Score 19; DB 9; Length 2064;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 caaggggtaggagaaag 196
|||||
Db 76 CAAGGGTAGGAGAAAG 58

RESULT 42
BC005333/c
LOCUS BC005333 2127 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, interferon gamma receptor 1, clone MGC:12420
IMAGE:3950528, mRNA, complete cds.
ACCESSION BC005333
VERSION BC005333.1 GI:13529118
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2127)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 16 Row: k Column: 12
This clone was selected for full length sequencing because it